NPS-OR-00-007

NAVAL POSTGRADUATE SCHOOL
Monterey, California

***J-STOCHWARS* and Beyond:**
Models for Force Motion and Interaction that Represent Uncertain Perception

by

Donald P. Gaver
Patricia A. Jacobs
Mark A. Youngren
Samuel H. Parry

September 2000

Prepared for: Conventional Forces Division, J8, the Joint Staff, The Pentagon, Washington DC;
DSO National Laboratories, Singapore; and
Institute for Joint Warfare Analysis

NAVAL POSTGRADUATE SCHOOL
MONTEREY, CA 93943-5000

Rear Admiral D. R. Ellison
Superintendent

Richard Elster
Provost

This report was prepared for and funded by the Conventional Forces Division, J8, the Joint Staff, The Pentagon, Washington DC; DSO National Laboratories, Singapore; and the Institute for Joint Warfare Analysis.

Reproduction of all or part of this report is authorized.

This report was prepared by:

Donald P. Gaver
DONALD P. GAVER
Distinguished Professor of Operations Research

Patricia A. Jacobs
PATRICIA A. JACOBS
Professor of Operations Research

Mark A. Youngren
MARK A. YOUNGREN
Analyst, MITRE Corporation

for
SAMUEL H. PARRY
Professor Emeritus of Operations Research

Reviewed by:

for
R. KEVIN WOOD
Associate Chairman for Research
Department of Operations Research

Richard E. Rosenthal
RICHARD E. ROSENTHAL
Chairman
Department of Operations Research

Released by:

DAVID W. NETZER
Associate Provost and Dean of Research

REPORT DOCUMENTATION PAGE			Form approved OMB No 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE September 2000	3. REPORT TYPE AND DATES COVERED Technical

4. TITLE AND SUBTITLE *J-STOCHWARS* and Beyond: Models for Force Motion and Interaction that Represent Uncertain Perception	5. FUNDING MIPR NO. DJAM00109
6. AUTHOR(S) Donald P. Gaver, Patricia A. Jacobs, Mark A. Youngren, Samuel H. Parry	
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Naval Postgraduate School Monterey, CA 93943	8. PERFORMING ORGANIZATION REPORT NUMBER NPS-OR-00-07
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) Conventional Forces Division, J8, the Joint Staff, The Pentagon, Washington DC; DSO National Laboratories, Singapore; and Institute for Joint Warfare Analysis	10. SPONSORING/MONITORING AGENCY REPORT NUMBER
11. SUPPLEMENTARY NOTES	
12a. DISTRIBUTION/AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.	12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This report records work carried out on a theater-level simulation modeling effort entitled *J-STOCHWARS* (*JS*). It is a connected set of working papers: minimal attempt has been made to polish the prose, and redundancies occur. There remain needed additions and modifications. These will be made as the authors adapt the ideas for use in planning models that are under current, and future military development. Work on the simulation model has been supplanted by an *analytical scoping model, Battlespace Information War (BAT/IW)*.

The emphasis of the report is upon recognition, characterization and exploitation and control of aspects of *uncertainty* in military operations. It emphasizes simple low-resolution representations of the products of generic imaging sensors. Later work will deal with other input, such as ELINT and COMINT. We provide ideas as to how such information products can be combined or fused, basically using a Bayesian and Gaussian/Normal-distribution theory format for computational and conceptual convenience. We investigate processes for probabilistically updating opponent course of action (COA) estimates.

Many of the computational ideas and algorithms have been incorporated in a computer program termed Joint Warfare Analysis Experimental Prototype (JWAEP). The latter runs on a SUN computer.

14. SUBJECT TERMS	15. NUMBER OF PAGES
sensor information; sensor fusion; course of action inference; C4ISR; inference for the number of binomial trials	133
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL

NSN 7540-01-280-5800

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std 239-18

J-STOCHWARS and Beyond: MODELS FOR FORCE MOTION AND INTERACTION THAT REPRESENT UNCERTAIN PERCEPTION

D. P. Gaver

P. A. Jacobs

M. A. Youngren

S. H. Parry

SUMMARY

This report records work carried out on a theater-level simulation modeling effort entitled ***J-STOCHWARS*** (***JS***). It is a connected set of working papers: minimal attempt has been made to polish the prose, and redundancies occur. There remain needed additions and modifications. These will be made as the authors adapt the ideas for use in planning models that are under current, and future military development. Work on the simulation model has been supplanted by an ***analytical scoping model, Battlespace Information War (BAT/IW)***.

The emphasis of the report is upon recognition, characterization and exploitation and control of aspects of *uncertainty* in military operations. It emphasizes simple low-resolution representations of the products of generic imaging sensors. Later work will deal with other input, such as ELINT and COMINT. We provide ideas as to how such information products can be combined or fused, basically using a Bayesian and Gaussian/Normal-distribution theory format for computational and conceptual convenience. We investigate processes for probabilistically updating opponent course of action (COA) estimates.

Many of the computational ideas and algorithms have been incorporated in a computer program termed Joint Warfare Analysis Experimental Prototype (JWAEP). The latter runs on a SUN computer.

1. Background

It is a truism to state that when several military forces jointly interact within a theater to oppose the actions of another they do so under conditions of substantial uncertainty as to the outcome. In fact, Carl von Clausewitz has said "War is the realm of uncertainty: three quarters of the factors on which action in war is based are wrapped in a fog of greater or lesser uncertainty. A sensitivity and discriminating judgment is called for; a skilled intelligence to scent out the truth." The purpose of this report is to list many of the sources of that uncertainty and to explore ways of describing their possible impact on a theater-level conflict, as the latter may develop over time. The style used is that of quantitative mathematical modeling, but many of the numbers (*parameters*) are at best tentative, as are the ways of representing their interaction (the *models*). Mathematical modeling requires specificity to obtain actual numerical "results", but in many situations these results must be treated as suggestive and tentative, and neither accepted nor discounted too hastily, without assessing both vices and virtues of the *modeling option* currently in use. Our approach will be to propose and give arguments for, and against, a variety of such options for many aspects of theater-level conflict problems without exclusively endorsing any one.

Our purpose is to develop models that have the "feel" of aspects of real campaigns, but that omit many details. They are "low resolution" by design. In particular, the optional approaches should allow the user-analyst to efficiently experience *alternative futures* that might plausibly follow from adopting certain doctrinal options.

This report is a record of work done on various aspects of theater level combat modeling; the emphasis is on simple perception modeling and not on attrition as this might be linked to perception. Work on the latter is reported elsewhere.

1.1 Sources of Uncertainty

Sources of uncertainty in military operations, and options and variations in the modeling thereof, are many. For a recent general review see Hughes (1994). Prominent among them is the actual, operationally relevant, status of various *physical aspects of the theater environment*, e.g. of the geographical terrain and weather, as influenced by season and time-of-day; others can be identified. Abbreviate designation of this class by PEUF for *physical environment uncertainty factors*. PEUF influence speed and predictability of force movement, visibility and hence detection and classification of force elements, fortification potential or the defensibility of specific locations, communications, equipment performance and reliability, ultimately even morale, all to varying degrees, many of which may be anticipated by an opponent, but with inevitable error.

Another prominent and related uncertainty source has to do with the trustworthiness of *intelligence*, this term generically referring minimally to knowledge of the present location, activity and current motion of both opponents' forces by their respective opponents. This knowledge, with its inevitable uncertainty, depends on the types, manner of utilization, and product of *information-gathering sensory assets*. These range from human scouts and informants (HUMINT) to electronic intelligence (ELINT) acquisition assets such as radars or sonars, either stationary or carried on various mobile platforms (manned or unmanned reconnaissance, aircraft or ground vehicles, satellites, etc.) that actively search certain aspects of terrain for opponent asset occupancy, but with unavoidable omissions and inaccuracies.

Additional sources of information, also susceptible to errors and omissions of acquisition and interpretation, including active opponent deception, are a variety of assets such as satellites, manned reconnaissance platforms, UAVs, and TUGVs. All of these ultimately supply *images* that must be interpreted. We group these in the category abbreviated as IMINT. In addition, there is the information to be gained from sensing the communication emissions of an opponent force and interpreting their meaning. This area is called COMINT, and is useful but also susceptible to errors of misinterpretation and delay. The uncertainty associated with (only!) the physical performance of information-gathering assets, e.g. its probability of detection of a specified target asset such as a tank (or tank company) or armored personnel carrier, its probability of correct classification given detection, its accuracy of location of the sensed asset (conditional on the sensor system operator's skill and the surrounding physical and operational conditions) can be abbreviated as PIAU, or *physical intelligence asset uncertainty* factors.

When actual combat occurs further sources of uncertainty must be confronted, namely the *physical capabilities of the weapons* brought to bear. Even if intended targets are well-located and weaponry is extremely accurate ("smart") and well-matched to target vulnerability there are ample opportunities for errors and equipment (and operator) failures that can lead to partially- successful mission accomplishment, and hence requirement for follow-up action. We lump such *physical weapon asset uncertainty* into a category abbreviated by PWAU. It is apparent that even smart weapons can be decoyed or otherwise duped, which adds a further element of uncertainty, and necessitates thoughtful battle damage assessment (BDA), in turn influenced by PIAU as above.

The above rough categorization of the major physical components of uncertainty that must be considered by both opponents and allies in theater level warfare is not exhaustive. There are other far more tenuous components that are, however, significantly affected by the stark physical influences mentioned above. A major component of the overall uncertainty must be that concerning the territorial objectives of, say, an aggressor into a neighboring area, and the willingness or resolve to attain those objectives at the expense of losses of assets and his/her own territory. It is far more difficult to credibly quantify the uncertain tenacity of an enemy force in the face of actual or prospective attrition than it is to model physically-based uncertainties associated with unopposed enemy advance rate through somewhat unknown, and unexpectedly variable, terrain while accounting somewhat approximately for that force's size and weapons and intelligence capabilities. Opponent resolve and tenacity have often been modeled by *thresholds* that limit attrition in case combat occurs, perhaps by withdrawing, but also by calling for reinforcements, which are then unavailable to another area or conflict in a different sub-theater. Note that the thresholds can be of various forms: they can reflect own force losses (or prospective losses), or, additionally, estimates of opponent force losses; the latter may well be "known" only roughly. Thresholds may also be expressed as current or projected *rates* or estimated rates (derivatives) of loss, cf. Helmbold (1971).

A generic scenario of Red and Blue forces moving towards a common objective in an invasion, occupation, protection scenario is described in the next section. A brief description of the physical theater in terms of physical nodes and arcs is given. Increasingly detailed models of force movement along arcs are presented in Sections 2 and 3. Models for detection and tracking of units are also introduced. Procedures for estimating each side's units' velocities, arrival time at the objective and unit size are proposed and studied. In Sections 4 – 6, procedures for estimating the number of assets on a node or arc based on binomial, beta-binomial and multinomial models for sensor observations are proposed and studied. In Section 7, procedures for estimating the number and types of military units on a node/arc based on estimates of the number of assets at that node and neighboring nodes are studied. Section 8 describes a formulation of each side's course of action (COA) in a conflict. Procedures for one side to estimate, using sensor observations, the probability the other side is following each COA are described.

The present report is a progress statement that documents ideas and concepts that may be useful in other modeling projects. Some, but not yet all, of these ideas have been realized in JWAEP (Joint Warfare Analysis Experimental Prototype), a demonstration software package described in the JWAEP User's Manual; cf. Youngren (1996).

2. Invasion-Occupation-Protection Dynamics: A Simple Prototypical Example

We propose for initial discussion a simplified prototype for many real situations: Iraq's invasion of, and expulsion from, Kuwait, or its possible invasion of other neighboring countries; North Korea's possible invasion of South Korea, and perhaps others on a smaller scale. It is stipulated to possess these elements:

(a) A *region,* abbreviated $\mathcal{N}$, that is in dispute, possibly because of its desirable resources (oil, resort areas and gambling casinos, bountiful production of, say, hog jowls or titanium).

(b) One (or more) *unfriendly invading forces*, termed Red, that intend to capture territory and exploit resources of $\mathcal{N}$. They must enter the region by crossing its boundaries. Some entry points have advantages, i.e. are less well-protected or defensible than others. Let $R(t)$ be the size of Red force(s) at t. $R(t)$ is actually a vector whose components are numbers of force types and their locations; $R(t)$ is *Red ground truth* at time t.

(c) A *protective or defensive force*, or forces, that also originates outside of $\mathcal{N}$; call it Blue (B). Its force size is $B(t)$. If an offensive move by R occurs, B responds in one of several ways. $B(t)$ is also a vector as is $R(t)$; it is *Blue ground truth* at t. There is dynamic interaction between $R(t)$ and $B(t)$.

(d) A *civilian* or *non-combatant indigenous population*; it may well be that the population must be split into segments sympathetic to R, $C_R(t)$ in number, and those sympathetic to B, $C_B(t)$, with total $C(t) = C_R(t) + C_B(t)$. Finer distinctions are always possible, in which case, the above state variable may be a vector. In many situations affiliations would change as combat progresses and population elements alter their perceptions as to whom they might wish to ally themselves with. Considerations for such a population can well be a major concern in operations-other-than-war (OOTW) or peacekeeping scenarios.

(e) An *indigenous military force* able to delay, but not repel, an unfriendly invading force, or forces. Let $I(t)$ generally denote the size of that force at time t; realistically it will be vector-valued so as to represent force components: infantry, armor, artillery, air, etc. There may also be indigenous forces that are allies of the invaders; let the number of these be $I_R(t)$. Those allied with B are $I_B(t)$ in number.

Of course the implications of the above terminology need not hold for all situations, i.e. that the region, $\mathcal{N}$, is in friendly (to B) hands initially and is threatened by "bad guys" and protected by "good guys". In fact, $\mathcal{N}$ may be initially in the hands of R, and its domination could come under attack by B. And the geography must be specified in enough detail to portray a region and the relevant surrounding territory. If desired, regions can be represented as a union of non-overlapping subregions.

Geography: Arcs and Nodes

A *theater*, that is the geographical entities relevant to the elements (a) – (e) above, is conveniently represented by a system of *nodes*: physical locations, fixed in space, and of importance either because they have symbolic value, e.g. are centers of governmental authority, *or* are suitable for defense by fortification or asset concealment because of topography, *or* are significant locations for military assets such as airfields, surface-to-air missile sites, (relatively) immobile command centers or possibly C3/I facilities. Nodes are viewed as interconnected by *arcs*: road systems or other routes for the transit of ground combatants. For some purposes it is useful to represent arcs as a sequence of discrete locations (Subarcs, or *compartments*) between which military units transition.

In a naval context it may often be convenient to specify geographical locations as being within particular *broadened arcs* or spatial compartments, shaped as squares or hexagons.

A primitive arc-node setup appears below:

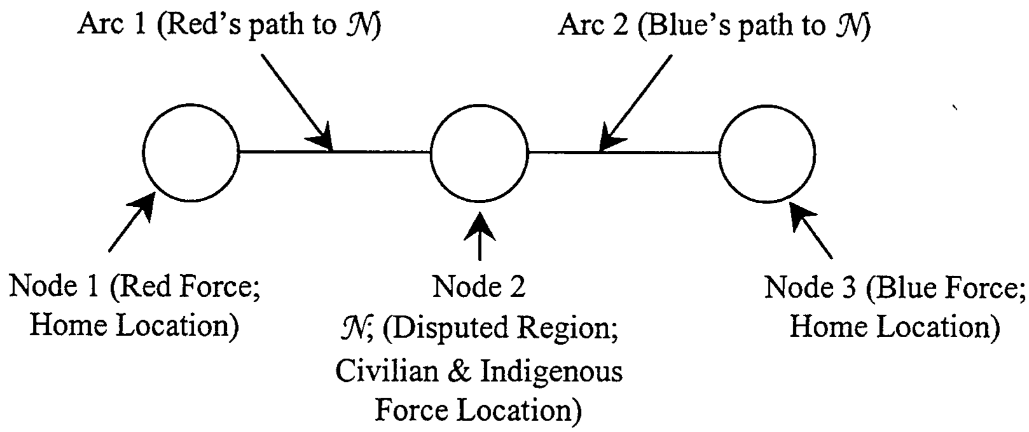


Figure 2.1

The above simplistic setup can be elaborated greatly: realistically there may be many routes from Node 1 to Node 2, and Node 3 to Node 2, possibly by way of intermediate nodes.

The arc-node representation is perhaps more appealing when alternative routes can be cleanly and crisply distinguished because of topography (a substantial ground force does not easily traverse mountain ranges) than when the topography is bland and neutral, as in a desert, or in the open ocean. From the point of view of air warfare, wherein aircraft from a source node may transit to a target area to conduct reconnaissance or an offensive strike, arcs are a less attractive representation than is a broad "tiling" of the region, possibly by square or hexagonal subregions or compartments. Sorties of aircraft simply move, in appropriate time, from tile or compartment (center) to a contiguous tile or compartment (center) as they transit a region. Alternatively, straight line flight paths could be scripted through the region.

Dynamics

Military operations by, say, Red, and in response by Blue, can be specified in terms of arcs and nodes. For example, in Figure 2.1 a Red force moves on Arc 1 from N1 to N2. Some options for the dynamics are listed below.

Option 1: Simple Random Time-On-Arc (Time-On-Node)

It is a reality that even an unopposed ground force unit will move across a given terrain route in a somewhat unpredictable, variable, or uncertain time. This is one consequence of PEUF (Section 1). The simplest way of representing one such transit time is to let it be a single random draw from an appropriate population, or realization, T, of a random variable, $\boldsymbol{T}$. Then if D is the distance, say from Node 1 to Node 2 in Figure 2.1, the velocity of transit is $V = D/T$, a realization of a random variable. Having obtained T it is then easy to locate the unit's location (center) at time $0 \leq t \leq T$ after it leaves the origin node if we assume a constant velocity: its distance from the starting node in the direction of motion is $D(t) = Dt/T$; this number locates the unit on the arc and is referred to as *ground truth*. Note that actual ground truth will vary across realizations of the model because of the random variability of T. Of course the distributional properties of $\boldsymbol{T}$ will potentially depend on terrain type, even direction of motion, (e.g. up or down hill), unit type, time of day, season of year, and perhaps other factors as well. Convenient candidate distributional forms are the Gaussian or normal (adjusted to truncate away possible negative values), the inverse Gaussian (which has naturally positive support), the lognormal, log-logistic, gamma, Weibull, or other distributions with positive support. A practical issue is that of actually specifying numerical parameters to specify the distributional form selected. Estimates of the mean and variance should be available from field data, but analyst or other expert judgment may well be required, at least initially. The effect of concomitant explanatory variables can be incorporated by regression modeling; see McCullagh and Nelder (1993). It will thus be possible to represent weather effects by regression.

The advantage of this approach is its simplicity. Disadvantages or unrealities include the fact that a unit's velocity across an arc is represented as constant once the time T is selected; of course this will change when another plausible history is selected, i.e. on another model evolution. If an event occurs while the unit is in transit, e.g. if it encounters an obstacle such as a mine field and waits to sweep it, or alternatively bypasses it at the cost of additional delay, a new random number must be drawn. Alternatively, the mean and variance of the original time, T, may be adjusted to reflect such extra delays. Also, if the unit in transit encounters an opponent and pauses for combat, another random draw will be required to represent the future motion of the unit. Note that for the purpose of reducing computing demand all elements or Subunits (e.g. companies) of a unit (e.g. brigade) are represented as moving together. They may be conveniently viewed as located in an assigned formation (template) within a moving rectangular matrix.

Option 2: Markovian Movement Through Subarcs

To better represent potential local in-transit variability of motion, suppose that an arc between two consecutive nodes is divided into several consecutive Subarcs or compartments; see below, where there are exactly k Subarcs.

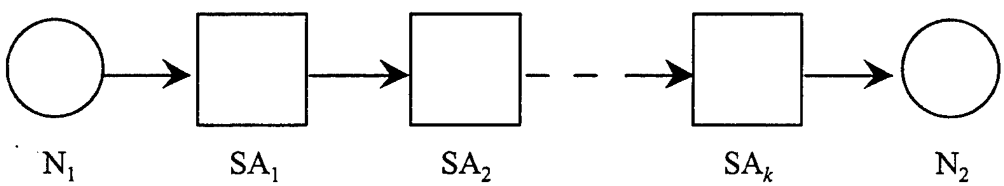


Figure 2.2

Imagine that a Unit (e.g. brigade) or collection of Subunits (e.g. companies composing a brigade) leave N_1 and enter SA_1 at $t = 0$; they remain in SA_1 for a random residence time, T_1, and thence move to SA_2, from which they transit to SA_3 after random residence time T_2, and so on through SA_k and to N_2. If the above process is defined in continuous time ($t \in R^+$) and the transit/ residence times are independently and exponentially distributed with parameters λ_j then the location of a unit, i.e. the Subarc occupied, at time t is a Markov process in continuous time. If the process is defined in discrete time, and the transit/residence times are independently and geometrically distributed with probability of one-stage advance being p_j (and staying in place on SA_j with probability $1 - p_j$) then the location of the unit at time ($t \in (0, 1, 2, \ldots)$ is Markov in discrete time.

Example: Suppose the motion process is in continuous time and all Subarc times, $\boldsymbol{T}_1$, $\boldsymbol{T}_2, \ldots$ are independently and identically exponentially distributed, with parameter λ. Then the unit moving is located at/within Subarc j ($j \leq k$) at time t with the *Poisson* probability

$$P\{\text{Unit in subarc } j \text{ at } t\} = e^{-\lambda t}\frac{(\lambda t)^j}{j!}; \tag{2.1}$$

the most likely Subarc to be in at t is $j = [\lambda t]$; if $[\lambda t] \geq k + 1$ the Unit has arrived at N_2. Now to calibrate parameters λ and k to a specified mean unit velocity $\overline{V}$ at which the distance D is covered we require

$$\frac{k}{\lambda} = \frac{D}{\overline{V}} = \overline{T}. \tag{2.2}$$

Suppose in addition we want to specify that the standard deviation of $\overline{T}$ is a fraction, f, of $\overline{T}$. This means that

$$\sqrt{\frac{k}{\lambda^2}} = f\overline{T} = f\frac{k}{\lambda} \tag{2.3}$$

which implies that $k = 1/f^2$. Thus if $f = 0.10$ the number of Subarcs $k = 100$, while if $f = 0.25$, $k = 16$: the larger the variability the smaller the number of Subarcs required. With a moderate number of Subarcs (even 16) it will usually be adequate to use the normal approximation to the Poisson. The parameter λ, or rate at which a Subunit moves off a Subarc is, from (2.2), obtained as $\lambda = k/\overline{T}$, or, more intuitively, $1/\lambda = \overline{T}/k$.

An extension of the previous simple model is to allow backward as well as forward motion, and also pauses on an arc of more than one basic exponential duration.

Numerical Illustration 2.1. Independent Simultaneous Approach; Race Towards a Common Objective.

Figures 1 – 4 show the results of a simulation of the simultaneous approaches of Blue and Red forces towards the common objective $\mathcal{N} = N_2$. For this illustration the two forces move *independently* of each other. The Blue force starts from a distance of $D_B = 200$ miles and moves at mean velocity $V_B = 15$ mph; Red from $D_R = 100$ miles, mean velocity 10 mph. Each distance is divided into 20 Subarcs. There were 1000 replications. A force is said to be late if it is not the first to arrive at $\mathcal{N}$. The resulting lateness time is the length of time after the first arrival until arrival of the late force. Of special tactical/operational interest is the *lateness* of either Blue or Red: for the present situation Blue's lateness probability is estimated to be about 0.80; if late, Blue's lateness time mean is 4.4 hours, but the distribution of lateness time, given Blue arrives last, is substantially spread out, providing Red considerable time to prepare a position on N_2. On the other hand, Red is late with probability about 0.20; if late, Red's lateness time mean is only about 2 hours and seldom more than 5 hours. Note that a deterministic assessment of arrival order would be based on comparison of $D_B/\overline{V}_B = \overline{T}_B = 200/15 = 13.3$ and $D_R/\overline{V}_R = \overline{T}_R = 100/10 = 10$, so on that basis Blue would *always* arrive 3.3 hours after Red. A more refined and easy calculation is based on a normal approximation

$$\text{Probability of Blue Lateness} = P\{T_B - T_R \geq 0\} \cong 1 - \Phi\left(\frac{\overline{T}_R - \overline{T}_B}{\sqrt{\sigma_R^2 + \sigma_B^2}}\right) \cong 0.81 \qquad (2.4)$$

where we have computed

$$\sigma_B^2 = n_B\left(\frac{D_B}{V_B n_B}\right)^2 = \frac{\overline{T}_B^2}{n_B} = 8.8$$

$$\sigma_R^2 = \frac{\overline{T}_R^2}{n_R} = 5.0. \qquad (2.5)$$

(If the simulated variances are used the result is 0.81, quite close to the simulation value.)

Discussion. The above illustration is quite rudimentary and unrealistic in that no sensor activities and resulting information, and subsequent action changes, are modeled. It does, however, illustrate the effect of variability in transit.

Numerical Illustration 2.2. Motion with Detection and Estimation of Arrival Time.

Consider this movement scenario, similar to the above, but modeling detection by Blue. Once again, Red is initially located a distance D_R from the objective $\mathcal{N}$, Blue is initially located a distance D_B from the objective $\mathcal{N}$, the mean velocity for Red (respectively Blue) is $\overline{V}_R$ (respectively $\overline{V}_B$).

Assume the distance D_R (respectively D_B) is divided into N_R (respectively N_B) equal Subarcs. Assume the time to transverse Subarc i for Red (respectively Blue) $T_R(i)$ (respectively $T_B(i)$) has a distribution with mean $(D_R/\overline{V}_R)/N_R$ (respectively $(D_B/\overline{V}_B)/N_B$).

Example. $D_R = 100$ miles $\quad D_B = 200$ miles

$\overline{V}_R = 10$ miles/hr. $\quad \overline{V}_B = 15$ mph

Both Blue and Red traverse 30 Subarcs. Suppose now that Blue can detect Red. The probability that he detects Red on a Subarc, given her presence there, is δ, temporarily assumed constant. Thus it may be representative of wide-area theater-level, but not unit-level, overhead assets, implicitly cueing a "soda straw" sensor such as UAV. The number of the Subarc on which first detection is made has a geometric distribution (if there were an unlimited number of subarcs)

$$P\{M = k\} = \delta(1-\delta)^{k-1} \qquad k = 1,2,\ldots.$$

Assume that Blue tracks Red for d_t Subarcs after detection at M. Blue's estimate of the velocity of Red is

$$\hat{V}_R = \left[\frac{\sum_{i=M+1}^{M+d_t} T_R(i)}{d_t\, D_R/N_R}\right]^{-1} \tag{2.6}$$

where $T_R(i)$ is a time to transit Subarc i, a realization of $\boldsymbol{T}_R(i)$. Blue's estimate of Red's time of arrival at $\mathcal{N}$ is

$$\hat{A}_R = D_R\left[1-\frac{M+d_t}{N_R}\right]\Big/\hat{V}_R = \frac{\left[N_R-(M+d_t)\right]}{d_t D_R}\sum_{i=M+1}^{M+d_t} T_R(i). \tag{2.7}$$

The elapsed time since the start of the replication is $\sum_{i=1}^{M+d_t} T_R(i)$. The number of the Subarc that Blue is on at this time is

$$\min\left\{k\colon \sum_{i=1}^{M+d_t} T_R(i) \le \sum_{i=1}^{k} T_B(i)\right\} \equiv N_E(B) \tag{2.8}$$

where $T_B(i)$ is Blue's transit time on his Subarc i. Given Blue's estimate of his own velocity, V_B, Blue can estimate an approximate probability that Red will arrive at $\mathcal{N}$ x time units before Blue will.

$$1-p_L(x) = P\left\{\sum_{i=N_E(B)+1}^{N_B} T_B(i) > \hat{A}_R + x\right\}. \tag{2.9}$$

If the $\boldsymbol{T}_B(i)$ are exponentially distributed then the above calculation involves the survivor function of a gamma random variable with shape parameter $N_B - N_E(B)$.

Blue can potentially use the above results to calculate the chance of successfully carrying out certain actions to slow or damage the Red force, e.g. before that force reaches $\mathcal{N}$; Blue may speed up, send strike or assign indirect fire assets to attrit and delay Red, or use assets such as FASCAM (Family of Scattered Mines; an air or artillery-delivered minefield) to create obstacles in Red's path.

Numerical Illustration 2.3. Example.

Figures 5 – 10 display results for a simulation with 1000 replications in which the time to transit a Subarc is normally distributed. The number of Subarcs is 20 for both Red and Blue. The mean time for Red to transit a Subarc is $(D_R/\overline{V}_R)/20$ with standard deviation $(0.3)(D_R/\overline{V}_R)/20$. The mean time for Blue to transit a Subarc is $(D_B/\overline{V}_B)/20$ with standard deviation $(0.3)(D_B/\overline{V}_B)/20$. Figures 5 and (respectively 6) show histograms of the amount of time to transit the arc for Red (respectively Blue).

The probability of Blue detecting Red on a Subarc is $\delta = 0.3$. Figure 7 displays the number of the Subarc on which Red is detected.

Blue then uses two additional Subarcs to estimate the velocity of Red. There are 3 replications in which Red is first detected on Subarcs 18 – 20. For these replications Blue does not estimate the velocity of Red. Figure 8 displays a histogram of Blue's estimates of the velocity of Red. Figure 9 displays a histogram of Blue's estimate of the mean time remaining until Red achieves his objective. Figure 10 displays a histogram of Blue's estimate of the probability he achieves the objective first.

Numerical Illustration 2.4. Velocity Estimation by Exponential Smoothing.

Figures 11 – 14 show the results of a simulation in which there are 1000 replications. In each replication Red is $D_R = 100$ mi. away from his objective and is traveling with a mean speed $\overline{V}_R = 10$ mph. The arc along which Red is traveling is evenly divided into 20 Subarcs; thus each Subarc has a distance $D_R(S) = 100/20 = 5$. The time Red takes to travel a Subarc is normal with mean $D_R/(\overline{V}_R \times 20) = 0.5$ hours and standard deviation 0.2×0.5 hours; the normal random numbers are left censored at 0.05. The probability Blue detects Red on a Subarc is 0.3. Once Blue detects Red he is able to track Red.

One procedure for Blue to use to estimate Red's velocity once Red is detected is exponential smoothing. Let $\hat{V}(i)$ be the estimate of Red's velocity after Red passes through Subarc i after Blue detects Red. Blue's estimate of Red's velocity after Subarc $i+1$ is

$$\hat{V}(i+1) = (1-\alpha)\hat{V}(i) + \alpha\left[\frac{T_R(i+1)}{D_R(S)}\right]^{-1}$$

where $T_R(i + 1)$ is Red's time to traverse Subarc $i + 1$. In the simulations $\alpha = 1/3$. The estimated velocity can be used to estimate Red's time of arrival at the objective.

Figure 11 displays the histograms of the estimate of Red's velocity. One histogram displays the estimate obtained by dividing the distance of two Subarcs by the time Red takes to traverse the two Subarcs after the Subarc on which he is detected. The other histogram displays the histogram of the exponential smooth estimate using those Subarcs after detection that were traversed before time 6; the time of 6 was chosen arbitrarily as a perception update time. Those replications for which there are not two Subarcs after detection available for estimation before time 6 are not used. As expected, the exponential smoothed estimate has less variability than the two Subarc estimates.

Figure 12 displays histograms of the error of the estimate of Red's time of arrival at his objective. The estimate using the exponential smoothed estimate of velocity is less variable than that for the velocity estimate using two Subarcs; this is to be expected.

Figures 13 – 14 display results using the exponential smooth estimator of velocity at times 3 and 6. In each case a replication is deleted if it does not have two Subarcs to use before reporting. Figure 13 shows the histograms of the velocity estimates and Figure 14 shows histograms of the errors of the estimates of Red's time of arrival at the objective. The additional observation time to estimate the velocity of Red not surprisingly decreases the variability of the estimates.

Example 2.5: Unit Markov Motion

Suppose, for specificity, that a Red Unit of population size R (measured in number of Red Subunits, e.g. companies if R is one or more regiments or brigades) enters Arc 1 between N_1 and N_2 (Figure 2.2) via SA_1 at $t = 0$: Let $R_1(0) = R$ and let $R_j(t) = R$ if the Unit is on Subarc j at time t; otherwise $R_j(t) = 0$. Represent the Unit's motion as probabilistic as follows:

$$P\{\text{Red Unit moves from SA}_j \text{ to SA}_{j+1} \text{ in time step } (t, t+dt) | X_R(t) = x\}$$
$$= r_j(t, dt; x)$$

otherwise it remains at SA_j for $(t, t + dt)$ with probability $1 - r_j(t, dt; x)$. The time increment dt need not be infinitesimal, e.g. it may be one hour.

The (vector-valued) variable $X_R(t)$ represents a general external influence, based on perception, upon the probability of advance by Red. There exists a corresponding external influence variable, $X_B(t)$, for Blue.

For instance $X_R(t)$ may reflect the perception of Red that a Blue force of some estimated size is concentrated at a corresponding Subarc of Arc 2 (Figure 2.1) *and* that the corresponding probability of advance to disputed Node 2 has a specified estimated current value; this might be such that the Blue force will reach N_2 (=$\mathcal{N}$) within a specified (small) number of time steps with high probability. If this results in Blue reaching N_2 sufficiently in advance of Red to fortify and hence obtain an advantage, Red must decide among operational options, such as (a) continue as planned; (b) speed up and reach N_2 first (with high probability) and itself fortify N_2; (c) place obstacles in Blue's path (e.g. emplace a minefield, or setup an ambush, at a Subarc between Blue's current location and N_2 so as to delay Blue's progress, simultaneously proceeding towards N_2; (d) other (many possibilities!). Blue has comparable operational options.

Example 2.6: Suppose there are $C_j(t)$ Subunits on/in Subarc j at time t; $j \leq k$ in Figure 2.2. Then a number, $E_j(t, t + dt)$, may be ordered to move (emigrate) from Subarc j to Subarc $j + 1$ in time dt (for the present dt is a fixed non-zero time interval analogous, but not necessarily equal, to the PRT (perception review time) Δ). Likewise, a number $E_{j-1}(t, t + dt)$ on $j-1$ may emigrate from Subarc $j-1$ to Subarc j, leaving a net number on Subarc j equal to $C_j(t + dt) = C_j(t) + E_{j-1}(t, t + dt) - E_j(t, t + dt)$. The above transfer of Subunits is only feasible if there is a positive number of Units or Subunits on Subarcs $j-1$

and j to immigrate and emigrate, which means that both E_{j-1} and E_j are functions of their parent Subarc population levels, $C_{j-1}(t)$ and $C_j(t)$, being zero if those population levels are zero.

Example (Continued): Suppose that for the j^{th} Subarc

$$E_j(t,t+dt)=C_j(t)\left(1-e^{-\xi_j dt}\right);\quad 1\le j\le k. \tag{2.10}$$

Then

$$\begin{aligned} C_j(t+dt) &= C_j(t)+C_{j-1}(t)\left(1-e^{\xi_{j-1}dt}\right)-C_j(t)\left(1-e^{-\xi_j dt}\right) \\ &= C_j(t)e^{-\xi_j dt}+C_{j-1}(t)\left(1-e^{-\xi_{j-1}dt}\right) \end{aligned} \tag{2.11}$$

which is feasible, giving a positive result. Furthermore it is readily advanced in time, e.g. starting from the initial condition $C_1(0)=C$, where C is the population size of a Unit about to move onto the arc at $t=0$.

$$\begin{aligned} C_k(t+2dt) &= C_k(t+dt)e^{-\xi_k dt}+C_{k-1}(t+dt)\left(1-e^{\xi_{k-1}dt}\right) \\ &= \left[C_k(t)e^{-\xi_k dt}+C_{k-1}(t)\left(1-e^{-\xi_{k-1}dt}\right)\right]e^{-\xi_k dt} \\ &\quad +\left[C_{k-1}(t)e^{-\xi_{k-1}dt}+C_{k-2}(t)\left(1-e^{-\xi_{k-2}dt}\right)\right]\left(1-e^{-\xi_{k-1}dt}\right) \\ &= C_k(t)e^{-2\xi_k dt}+C_{k-1}(t)\left[\left(1-e^{-\xi_{k-1}dt}\right)e^{-\xi_k dt}+\left(1-e^{-\xi_{k-1}dt}\right)e^{-\xi_{k-1}dt}\right] \\ &\quad +C_{k-2}(t)\left(1-e^{-\xi_{k-2}dt}\right)\left(1-e^{-\xi_{k-1}dt}\right). \end{aligned} \tag{2.12}$$

Example (Continued): If $\xi_k=\xi$, i.e. if the transfer probability is the same for all Subarcs it can be seen that, putting $ndt=t$,

$$C_k(t)\equiv C_k(ndt)=\sum_{j=0}^{k}C_j(0)\binom{n}{k-j}\left(1-e^{-\xi dt}\right)^{k-j}\left(e^{-\xi dt}\right)^{n-(k-j)}. \tag{2.13}$$

In words, if there are initially $C_j(0)$ Subunits on/in Subarc j $(0\le j\le k)$ then any one of these must make exactly $k-j$ transitions between intervening Subarcs in n time steps in order to be on/in Subarc $k \geq j$ at time ndt; there are $\binom{n}{k-j}$ ways of selecting those that "make it". The formula (2.7) shows that if $C_0(0) = C$ Subunits start at time 0 then the number at exactly the k^{th} Subarc at time $ndt = t$ is

$$C_k(t) = C\binom{n}{k}\left(1 - e^{-\xi dt}\right)^k \left(e^{-\xi dt}\right)^{n-k}. \tag{2.14}$$

The Subarc containing the maximum number of Subunits is approximately $k = n(1 - e^{-\xi dt})$ when n is large; the number of Subunits on that particular Subarc approaches zero as n becomes large, at a rate proportional to $1/\sqrt{n}$, and those nearby have about the same number: Subunits become quite dispersed.

Comment. The above transition law, with ξ or ξ_k constant, could be appropriate for a disorganized force advance — or perhaps better, retreat — but is not truly representative of the motion of a structured force's coordinated advance towards a territorial objective. However, the dispersive effect noted above can be remedied by introducing appropriate time dependence into the parameters ξ_k.

3. A Red-Transit, Blue-Perception Vignette

At time $t = 0$, a Red Unit (e.g. Army brigade), composed of C Subunits, moves off a physical node onto a transit node or arc. It spends a random transit time, $\boldsymbol{T}$, on the arc, after which it reaches a subsequent physical node. Blue's objective is to estimate current force size, C, on the arc from observations available from a sensor system.

Opponent (Blue) Perception: Observational Modeling

Suppose a Blue Overhead Sensor Suite is constantly observing the arc on which the Red Subunit is moving, and that the random time, $\boldsymbol{D}$, to *detect* any one of the Red Unit's Subunit components that leaves in a Lost or untracked state has exponential distribution with mean $1/\delta$ (rate δ). Having detected such it *tracks* that Subunit for a random time $\boldsymbol{L}$ until *loss of track* occurs; $\boldsymbol{L}$ is assumed exponential with mean $1/\eta$. Search resumes, and the unit is detected again at a time that is an independent replica of $\boldsymbol{D}$ (presuming it is still on the arc); it is again susceptible to loss, and so on. The alternating sequences $\{\boldsymbol{D}_k\}$ and $\{\boldsymbol{L}_k\}$ are assumed independent. Furthermore, individual Subunits are assumed to be detected independently. It is clear that the appearance on the arc of the C Subunits triggers C independent 2-state Markov chains. Note that a Subunit can be in a detection state at $t = 0$, as it begins transit, although the implication of the above discussion is that some Subunits begin their transits anonymously, i.e. in a Lost or undetected state. It follows that the number of units that start transit in a Lost state, $C_L(0)$, and those that start in a Detected/Tracked state, $C_D(0)$, are, at Blue's (first) Perception Review Time (PRT) $t = \Delta$, in either a Lost or Detected/Tracked state $\boldsymbol{C}_{LD}(\Delta)$, $\boldsymbol{C}_{LL}(\Delta)$, $\boldsymbol{C}_{DD}(\Delta)$, $\boldsymbol{C}_{DL}(\Delta)$, obeying the dynamics of a simple Markov chain with transition probabilities

$$P\{\boldsymbol{C}_{LD}(\Delta) = v_{LD}(\Delta)\} = \binom{C_L(0)}{v_{LD}(\Delta)} p_{LD}(\Delta)^{v_{LD}(\Delta)} \left(1 - p_{LD}(\Delta)\right)^{C_L(0) - v_{LD}(\Delta)} \tag{3.1}$$

$$P\{\boldsymbol{C}_{LL}(\Delta) = v_{LL}(\Delta)\} = \binom{C_L(0)}{v_{LL}(\Delta)} p_{LL}(\Delta)^{v_{LL}(\Delta)} \left(1 - p_{LL}(\Delta)\right)^{C_L(0) - v_{LL}(\Delta)} \tag{3.2}$$

$$P\{\boldsymbol{C}_{DD}(\Delta) = v_{DD}(\Delta)\} = \binom{C_D(0)}{v_{DD}(\Delta)} p_{DD}(\Delta)^{v_{DD}(\Delta)} \left(1 - p_{DD}(\Delta)\right)^{C_D(0) - v_{DD}(\Delta)} \tag{3.3}$$

$$P\{\boldsymbol{C}_{DL}(\Delta) = v_{DL}(\Delta)\} = \binom{C_D(0)}{v_{DL}(\Delta)} p_{DL}(\Delta)^{v_{DL}(\Delta)} \left(1 - p_{DL}(\Delta)\right)^{C_D(0) - v_{DL}(\Delta)} \tag{3.4}$$

where

$$\begin{aligned} p_{DD}(\Delta) &= \frac{\delta}{\eta+\delta} + \frac{\eta}{\eta+\delta}\, e^{-(\eta+\delta)\Delta} = 1 - p_{DL}(\Delta) \\ p_{LD}(\Delta) &= \frac{\delta}{\eta+\delta}\left(1 - e^{-(\eta+\delta)\Delta}\right) = 1 - p_{LL}(\Delta). \end{aligned} \tag{3.5}$$

Recall that in Ground Truth C Red Subunits are hypothesized to be in motion on the arc for time $\boldsymbol{T}$; assume for now that $\boldsymbol{T}$ is long enough so that departure from the arc does not occur. These same transition probabilities prevail at every future PRT, i.e. at $k\Delta$, $k = 1, 2, 3, \ldots$, so long as the Subunit remains on the arc.

It becomes necessary to specify Blue's perception at $t = 0$ of the Red force size actually departing. We state some options for both observations and inferences. Not all are totally satisfactory.

Option 1: Simple Moment Estimators

It is easy to see from (3.1) – (3.4) that if C units transit

$$\begin{aligned} E\left[C_D(\Delta)\middle|C_D(0)\right] &= C_D(0)p_{DD}(\Delta) + C_L(0)p_{LD}(\Delta) \\ &= C_D(0)p_{DD}(\Delta) + \left[C - C_D(0)\right]p_{LD}(\Delta). \end{aligned} \tag{3.6}$$

Thus if $\boldsymbol{C_D}(\Delta)$ is observed, giving $C_D(\Delta)$, a straightforward candidate moment estimate for C, is based on only the first two observations

$$\hat{C}(\Delta)^{\#} = \frac{C_D(\Delta) + C_D(0)\left[p_{LD}(\Delta) - p_{DD}(\Delta)\right]}{p_{LD}(\Delta)}; \tag{3.7}$$

by stationarity an estimate based on any two consecutive observations is

$$\hat{C}(k\Delta)^{\#} = \frac{C_D(k\Delta) + C_D((k-1)\Delta)\left[p_{LD}(\Delta) - p_{DD}(\Delta)\right]}{p_{LD}(\Delta)}. \tag{3.8}$$

Unfortunately these estimates, while technically unbiased, may assume negative values for certain observational values. To avoid this inadmissibility they may be modified to

$$\hat{C}(k\Delta) = \max\left[\hat{C}(k\Delta)^{\#}, 0\right] \tag{3.9}$$

which of course makes the estimate non-linear and hence awkward to study mathematically.

Time-Averaged Moment Estimates

It is intuitively clear that, so long as Subunits remain on the arc, an improved estimate should result from an appropriate averaging process. One such is to simply average all past estimates with equal weights: if PRT is currently $j\Delta$ then

$$\begin{aligned}\overline{C}(j\Delta) &= \frac{1}{j}\sum_{k=1}^{j}\hat{C}(k\Delta) \\ &= \frac{1}{j}\hat{C}(j\Delta) + \frac{j-1}{j}\overline{C}((j-1)\Delta)\end{aligned} \tag{3.10}$$

the latter recursive update formula is convenient.

An attractive alternative is the *exponentially weighted moving average* (EWMA):

$$\widetilde{C}(j\Delta) = \alpha\hat{C}(j\Delta) + (1-\alpha)\widetilde{C}((j-1)\Delta), \quad 0 < \alpha < 1. \tag{3.11}$$

This estimate, e.g. with $\alpha = 1/3$, automatically and mechanically reduces dependence upon the past in a way that the unweighted average, $\overline{C}$, does not. It does so without regard to the relationship of the most current estimate, $\hat{C}(j\Delta)$, to the immediate past, so it may be slow to adapt to events such as abrupt changes in arc occupancy, or to trends in time (adaptation comes by modifying α as a function of recent radical changes in $\hat{C}(j\Delta)$; details remain for new work). But its simplicity and computational convenience cannot be denied or improved upon.

Utilization of the Estimate in Model Context

Within an evolution of the model it will be a simulated sequence of estimated C-values, i.e. either $\overline{C}(j\Delta)$ or $\widetilde{C}(j\Delta)$, $j = 1, 2, \ldots$, that will be presented to the analyst who guides one side's actions. Corresponding estimates must be introduced into algorithms that govern the opponent's behavior. These present estimates may be augmented by information as to the likely range of force-size (C-values) consistent with the estimates, and by further information concerning Subunit types; all such information elements or perception components are explicitly modeled as afflicted by (realistic) error to some degree. The analyst will then make decisions concerning future moves, e.g. by Blue, in the face of the perceptions so delivered.

Option 2: (Quasi) Likelihood Estimation

Likelihood methodology is an alternative approach to combining information from previous observations. The awkward form of the likelihood function associated with the current Markov model for $C_D(k\Delta)$ leads to use of a moment-matched Normal approximation or *quasi likelihood*, see Nelder and McCullagh (1983). Assume then that $C(k\Delta)$ has approximate conditional density

$$f(c_k;c_{k-1};C)=\frac{\exp\left\{-\frac{1}{2}\left[c_k-\left(c_{k-1}p_{DD}+(C-c_{k-1})p_{LD}\right)\right]^2\Big/\sigma_k^2\right\}}{\sqrt{2\pi\sigma_k^2}} \tag{3.12a}$$

where

$$\sigma_k^2=c_{k-1}p_{DD}(1-p_{DD})+(C-c_{k-1})p_{LD}(1-p_{LD}) \tag{3.12b}$$

and c_k and c_{k-1} are the observed counts at times $k\Delta$ and $(k-1)\Delta$ respectively. Now the likelihood function for unknown (but presumed constant) C is

$$L(C,j;\text{data})=\prod_{k=1}^{j}f(c_k;c_{k-1},C). \tag{3.13}$$

Rearrangement puts this in the form

$$\begin{aligned}L(C,j;\text{data})&=\prod_{k=1}^{j}\frac{e^{-\frac{1}{2}\left[c_k-c_{k-1}(p_{DD}-p_{LD})-Cp_{LD}\right]^2/\sigma_{k-1}^2}}{\sqrt{2\pi\sigma_{k-1}^2}}\\&=\frac{e^{-\frac{1}{2}\left[c_j-c_{j-1}(p_{DD}-p_{LD})-Cp_{LD}\right]^2/\sigma_j^2}}{\sqrt{2\pi\sigma_j^2}}\prod_{k=1}^{j-1}\frac{e^{-\frac{1}{2}\left[c_k-c_{k-1}(p_{DD}-p_{LD})-Cp_{LD}\right]^2/\sigma_{k-1}^2}}{\sqrt{2\pi\sigma_{k-1}^2}}.\end{aligned} \tag{3.14}$$

Of course σ_k^2 depends upon the unknown C, but also upon all previous observations. We propose to iteratively update the approximate likelihood estimate of C as follows:

(a) write the likelihood up to $(j-1)\Delta$ in the normal form

$$L(C, j-1; \text{data}) = \frac{e^{-\frac{1}{2}[C-\mu_{j-1}]^2/\tau_{j-1}^2}}{\sqrt{2\pi\tau_{j-1}^2}}. \tag{3.15}$$

This replaces the product to $j-1$ on the right-hand side of (3.14);

(b) complete the square to calculate the likelihood up to $j\Delta$ in the form

$$L(C, j; \text{data}) = \frac{e^{-\frac{1}{2}[C-\mu_j]/\tau_j^2}}{\sqrt{2\pi\tau_j^2}} \tag{3.16}$$

where

$$\mu_j = \frac{\left(\mu_{j-1}/\tau_{j-1}^2\right) + \left[\left(c_j - c_{j-1}(p_{DD} - p_{LD})\right)/p_{LD}\right]/\left(\tilde{\sigma}_j^2/p_{LD}^2\right)}{1/\tau_{j-1}^2 \; + \; 1/\left(\tilde{\sigma}_j^2/p_{LD}^2\right)} \tag{3.17a}$$

and

$$\tau_j^2 = 1/\left(1/\tau_{j-1}^2 \; + \; 1/\left(\tilde{\sigma}_j^2/p_{LD}^2\right)\right) \tag{3.17b}$$

where $\tilde{\sigma}_j^2 = c_{j-1}p_{DD}(1-p_{DD}) + \max\{(\mu_{j-1} - c_{j-1})p_{LD}(1-p_{LD}), 0\}$.

Since σ_{j-1}^2 depends upon the unknown C-value, replace that value by μ_{j-1}, available at $(j-1)\Delta$; actually it is best to apply the formula (3.12b) altered to replace $C - c_{j-1}$ by $\max(\mu_{j-1} - c_{j-1}, 0)$. This permits a numerical value for μ_j to be calculated; clearly this process may be iterated.

(c) The above expresses uncertainty about C in a Bayesian style, effectively attributing to C a density with mean μ_j and variance τ_j^2 at time $j\Delta$. Observe that μ_j actually is, as written in (3.17a), a linear combination or weighted average of the most recent moment estimator, $\hat{C}_D(j\Delta)$, as in (3.8), and the most recent point estimate, μ_{j-1}. The form is exactly analogous to the exponentially weighted moving average but with the weights determined in a systematic manner, i.e. updated by use of the estimated variance of the observations and the (approximate) model. Additionally there is the build-in assessment of the variability of the estimate furnished by τ_j^2.

(d) The above setup can be made adaptive to changes, e.g. of arc content, by *discounting the older likelihood*. This leads to scaling up τ_{j-1}^2 by a factor $K_d > 1$: in (3.17) simply replace τ_{j-1}^2 by $K_d \tau_{j-1}^2$. Clearly more weight is now placed on the current estimate.

Option 3: Maximum Quasi-Likelihood and Laplacianized Quasi-Likelihood

An alternative to the simple iterative procedure of Option 2 is to collect all of the C-dependent terms of the quasi likelihood function (3.12) into the exponent:

$$Q(C) =$$

$$-\frac{1}{2}\left[c_k - \left(c_{k-1}p_{DD} + \left(C - c_{k-1}\right)p_{LD}\right)\right]^2 \frac{1}{c_{k-1}p_{DD}\left(1-p_{DD}\right) + \left(C-c_{k-1}\right)p_{LD}\left(1-p_{LD}\right)} \tag{3.18}$$

$$-\frac{1}{2}\ln\left[c_{k-1}p_{DD}\left(1-p_{DD}\right) + \left(C - c_{k-1}\right)p_{LD}\left(1-p_{LD}\right)\right].$$

Now minimize Q with respect to C so as to maximize the likelihood: differentiate with respect to C, set the derivative equal to zero and solve the resulting quadratic for the quasi mle; tedious algebra gives the result

$$\hat{C}_{ML}(k) = \frac{-b(k) + \left\{\max\left(\left(b(k)^2 - 4a(k)d(k)\right), 0\right)\right\}^{\frac{1}{2}}}{2a(k)} \tag{3.19}$$

where

$$a(k) = p_{LL}\left(1-p_{LL}\right)^3$$

$$b(k) = \left[p_{LL}\left(1-p_{LL}\right)\right]^2 + 2\left(1-p_{LL}\right)^2 c_{k-1}p_{DD}\left(1-p_{DD}\right)$$

$$x(k) = c_k - c_{k-1}\left[p_{DD} - \left(1-p_{LL}\right)\right]$$

$$d(k) = -\left[2\left(1-p_{LL}\right)c_{k-1}p_{DD}\left(1-p_{DD}\right)x(k)\right.$$

$$\left. + p_{LL}\left(1-p_{LL}\right)x(k)^2 - p_{LL}\left(1-p_{LL}\right)c_{k-1}p_{DD}\left(1-p_{DD}\right)\right].$$

The second derivative can be calculated and hence the approximate variance is

$$\sigma_{ML}^2(k) = v(k)^2 \times \left[\left(1-p_{LL}\right)^2 v(k) + 0.5\left[p_{LL}\left(1-p_{LL}\right)\right]^2\right]^{-1} \tag{3.20}$$

where $\nu(k) = p_{LL}(1-p_{LL})\hat{C}_{ML}(k) + c_{k-1}p_{DD}(1-p_{DD})$.

If $\hat{C}_{ML}(k)$ is now treated as normally distributed with the above parameters $\hat{C}_{ML}(k)$ and $\sigma^2_{ML}(k)$ its contribution can be incorporated into an updating procedure analogous to that of (3.14) – (3.17).

An alternative to the above procedure is to treat the quasi likelihood as a density for ***C***, treated as a random variable à la Bayes. Rewrite (3.12) as

$$g(C;\text{data}) = K\frac{e^{-\frac{1}{2}(C-\hat{C})/(aC+b)}}{\sqrt{2\pi}\sqrt{aC+b}} \tag{3.21}$$

where $\hat{C}$ is an abbreviation for $\hat{C}_D(\Delta)^{\#}$ or $\hat{C}_D(k\Delta)^{\#}$ as in (3.7), and a and b depend on data c_k and c_{k-1}, and parameters p_{DD} and p_{LD}.

(a) normalize, by determining K so that the approximate density

$$K\int_0^\infty \frac{e^{-\frac{1}{2}(C-\hat{C})^2/(aC+b)}}{\sqrt{2\pi}\sqrt{aC+b}}\,dC = 1 \tag{3.22}$$

approximately, using Laplace's method; cf. Tierney and Kadane (1986); do so by writing the integrand as $e^{Q(C)}$, with

$$Q(C) = -\frac{1}{2}(C-\hat{C})^2/(aC+b) - \frac{1}{2}\ln(aC+b) \tag{3.23}$$

and determine the minimizing value of $\hat{C}$ and the second derivative at that point; this leads to

$$K\int_0^\infty \frac{e^{-\frac{1}{2}(C-\hat{C}_{ML})^2/\sigma^2_{ML}}}{\sqrt{2\pi}}\,dC \cong 1 \tag{3.24}$$

from the previous development.

$$K \sim \left(1/\sqrt{\sigma^2_{ML}}\right)e^{-Q(\hat{C}_{ML})}.$$

Similarly,

$$E(aC+b)=\int_0^\infty (aC+b)g(C;\text{data})dC \tag{3.25}$$

can be approximated using Laplace's method by writing the integrand as a $Q_m(C)$ with

$$Q_m(C)=-\frac{1}{2}(C-C_m)^2/(aC+b)+\frac{1}{2}\ln(aC+b) \tag{3.26}$$

and determine the minimizing value of C_m, $\hat{C}_m$ and the second derivative at that point. This leads to

$$\begin{aligned} E(aC+b) &\approx Ke^{Q_m(\hat{C}_m)}\sqrt{\sigma_m^2}=e^{Q_m(\hat{C}_m)-Q(\hat{C}_{ML})}\sqrt{\sigma_m^2/\sigma_{ML}^2} \\ &\equiv L_m \end{aligned} \tag{3.27}$$

where

$$\sigma_m^2=-\left[\frac{\partial^2}{\partial C^2}Q_m(\hat{C}_m)\right]^{-1}. \tag{3.28}$$

In a similar vein

$$E\left[(aC+b)^2\right]=\int_0^\infty (aC+b)^2 g(C;\text{data})dC \tag{3.29}$$

which can be approximated by writing the integrand as $e^{Q_V(C)}$ with

$$Q_V(C)=-\frac{1}{2}(C-C_V)^2/(aC+b)+\frac{3}{2}\ln(aC+b) \tag{3.30}$$

and determine the minimizing value of C_V, $\hat{C}_V$, and the second derivative at that point. This leads to

$$\begin{aligned} E\left[(aC+b)^2\right] &\approx Ke^{Q_V(\hat{C}_V)}\sqrt{\sigma_V^2}=e^{Q_V(\hat{C}_V)-Q(\hat{C}_{ML})}\sqrt{\sigma_V^2/\sigma_{ML}^2} \\ &\equiv L_V \end{aligned} \tag{3.31}$$

where

$$\sigma_V^2 = -\left[\frac{\partial^2}{\partial C^2} Q_V\left(\hat{C}_V\right)\right]^{-1}. \tag{3.32}$$

An estimate for C can be obtained from (3.27)

$$\hat{C}_L = \frac{L_m - b}{a}. \tag{3.33}$$

An estimate for the variance of C can be obtained from (3.31) as follows.

$$L_V \approx E\left[(aC+b)^2\right] = a^2 E\left[C^2\right] + 2abE[C] + b^2.$$

Thus,

$$\begin{aligned} E\left[C^2\right] &\approx \frac{L_V - 2ab\hat{C}_L - b^2}{a^2} \\ Var[C] &\approx \left(\left[L_V - 2ab\hat{C}_L - b^2\right]/a^2\right) - \hat{C}_L^2 \end{aligned} \tag{3.34}$$

If $\hat{C}_L$ is now treated as normally distributed with the above parameters $\hat{C}_L$ and variance Var[C] its contribution can be incorporated into an updating procedure analogous to that of (3.17).

Option 4: Numerical Integration of a Posterior for *C* Based on Quasi-Likelihood

An alternative that is of strongly Bayesian flavor begins with the quasi-likelihood expressed as a density for $\boldsymbol{C}$, i.e. (3.21), and numerically integrates to find $E[\boldsymbol{C}|c_k, c_{k-1}]$ and $Var[\boldsymbol{C}|c_k, c_{k-1}]$. Following this, it performs the iterative update à la (3.17). These are the steps:

(a) normalize (3.21), i.e. determine K so that

$$K\int_0^\infty \frac{e^{-\frac{1}{2}\left(x-\hat{C}\right)^2/(ax+b)}}{\sqrt{2\pi(ax+b)}}\ 1(ax+b)dx = 1; \tag{3.35}$$

where $1(ax + b) = 1$ if $ax + b > 0$ and is 0 otherwise;

(b) numerically integrate to obtain

$$E\left[\boldsymbol{C}\middle|c_k, c_{k-1}\right] = K\int_0^\infty x \frac{e^{-\frac{1}{2}(x-\hat{C})^2/(ax+b)}}{\sqrt{2\pi(ax+b)}}\ 1(ax+b)dx\,; \tag{3.36}$$

and

(c) numerically integrate to obtain

$$E\left[\boldsymbol{C}^2\middle|c_k, c_{k-1}\right] = K\int_0^\infty x^2 \frac{e^{-\frac{1}{2}(x-\hat{C})^2/(ax+b)}}{\sqrt{2\pi(ax+b)}}\ 1(ax+b)dx\,. \tag{3.37}$$

(d) Convert the latter *numbers* into a value for the $Var[\boldsymbol{C}]$.

(e) Apply (3.17) to update.

4. Sensor Models Revisited: Binomial, Multinomial, and Generalized

This section reports models for describing the output of *imaging sensors* in an operational context. They are applied in certain simulation (Monte Carlo) models (JSTOCHWARS/JWAEPS) to represent detection and (mis) classification of entities (enemy vehicles) on the surface of the earth from points above it, where these observation points can either be fixed in place (hovering) or in motion.

Basic Idea

A sensor's view of the earth surface is limited to a spatial region thereon; call this the *footprint*; in *JS* this is over an arc, or a node, but in reality is a (two-dimensional, sometimes more) region. The footprint can be virtually stationary, or can move in search of new opportunities. An image of the footprint elements detected by the sensor is portrayed on a screen. It is initially assumed here that, if there are C (generic) items (potential targets) on the surface (such as tanks, other wheeled vehicles, trucks) detectable by the sensor system, i.e. in *motion* for some (e.g. J-STARS), then these are *seen* on a glimpse or scan with *conditional probability* p (and not seen with probability $q = 1 - p$);

the events of seeing on successive scans are conditionally independent. Alternatively, a picture is on the screen, but the scan or glimpse is a time of residence of an operator foveal image over a subregion of the screen. Items within that subregion are candidates for operator attention (mis)identification, and communication towards weaponeers. The physical conditions that affect the probability of detection are (a) terrain, or sea state, including vegetative coverage, (b) the (radar) cross-section of vehicle types of interest; same for infrared, heat detection, (c) state of motion, or hiding, (d) *action* by enemy (Red) entities, such as combat/attack by them that facilitate detection, such as SAM site radiation during air defense (AD) activities.

Models

An initial model is that $\boldsymbol{R}$, the random number of entities (e.g. potential targets) *revealed* or *detected* on one sensor scan/glimpse is binomially distributed, conditionally on environmental factors. Likelihood and moment estimators *approximately* provide this simple estimator of C,

$$\hat{C} = \frac{\boldsymbol{R}}{p(\bullet)} \tag{4.1}$$

where $\bullet$ signifies explanatory variables. A candidate parametric model for this dependency is the *random hazard* (or variations thereof, to be described); see Gaver (1963),

$$p(\underline{v}, \varepsilon) = e^{-\varepsilon(-\ln p_0(\underline{v}))} \tag{4.2}$$

where $\ln p_0(\underline{v}) = \sum_{i=0}^{I} \beta_i v_i$, a regression term.

The vector variable $v = (v_1, v_2, \ldots, v_p)$ is one of observable explanatory variables such as range from sensor to target region, terrain type, cross-section of targets, atmospheric properties, etc. The (possibly vector) variable $\varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m)$ describes *random* environmental fluctuations; these may be common to several, e.g. consecutive, sensor scans because of cloud cover. The coefficient vectors $\underline{\beta}$ can be estimated from data, and modified for sensitivity checking purposes. The idea is that a sensor output is modeled as the outcome of C biased (not necessarily ½ – ½) but independent coin flips with probability of success simultaneously influenced by specified explanatory variables. The advantage of such a regression model is that it can be fitted using observations made under different conditions, i.e. on different target types, with different ranges, terrain types, i.e. by pooling data to estimate the regression coefficients $\underline{\beta}$. There are standard computer programs for carrying out maximum likelihood calculations. Then the expression, e.g. (4.3) can be used to predict the actual number of targets present.

Notice that the estimate now has the form

$$\hat{C}(\underline{v},\varepsilon) = \frac{\boldsymbol{R}}{p(\underline{v},\varepsilon)}. \tag{4.3}$$

This is conditional on ε; so this model has the effect of increasing the *observed* number, R, to account for those targets not observed; of course this adjustment is model-dependent.

Since, conditional on ε,

$$E[\boldsymbol{R} \mid \varepsilon] = C\, p(\underline{v}, \varepsilon), \tag{4.4}$$

the unconditional expectation of $\boldsymbol{R}$ is

$$E\{E[\boldsymbol{R} \mid \varepsilon]\} = C\, E[\, p(\underline{v}, \varepsilon)]\,; \tag{4.5}$$

so, if we assume the distribution of $p(\boldsymbol{v}, \varepsilon)$, given the explanatory variable values, $\underline{v}$, is well-enough known, an unbiased (approximately!) estimate of C is

$$\hat{\boldsymbol{C}} = \frac{\boldsymbol{R}}{E\left[p(\underline{v},\varepsilon)\right]}. \tag{4.6}$$

To obtain a numerical estimate, having observed $\boldsymbol{R}$ at t, $\boldsymbol{R}_t = r_t$, we write

$$\hat{c}_t = \frac{r_t}{E[p(\underline{v}_t, \varepsilon_t)]} \tag{4.7}$$

where it is being assumed for the present that $\{\varepsilon_t\}$ is a sequence of independent random variables with "known" distribution.

Note that we are allowing here (modeling) a situation in which all observers are simultaneously affected by a chance event, such as cloud cover. We are not (yet) representing adaptive action by the sensor-carrying platform, such as flying below the clouds (not always possible), or using different sensor technology, e.g. IR.

Variance

The estimate at time t, given by (4.7), is an instance of a random variable, (4.6). Conditional on ε_t, and using the binomial model,

$$E\left[\hat{C}_t \middle| \varepsilon_t\right] = \frac{E\left[\boldsymbol{R}_t \middle| \varepsilon_t\right]}{E[p(\underline{v}_t, \varepsilon_t)]} = \frac{C\,p(\underline{v}_t, \varepsilon_t)}{E[p(\underline{v}_t, \varepsilon_t)]} \tag{4.8}$$

whose expectation is the true value, C.

Consider the variance of $\hat{C}_t$, a measure of the variability of $\hat{C}_t$ around C, the true value. Note that this assumes just one glimpse "at C" while a given ε_t is present, if ε_t remains the same for several glimpses the model must reflect this, and can. First,

$$Var\left[\hat{C}_t \middle| \varepsilon_t\right] = \frac{C\,p(\underline{v}_t, \varepsilon_t)(1 - p(\underline{v}_t, \varepsilon_t))}{\left(E[p(\underline{v}_t, \varepsilon_t)]\right)^2}.$$

It is known that, unconditionally,

$$\begin{aligned} Var\left[\hat{C}_t\right] &= E\left\{Var\left[\hat{C}_t \middle| \varepsilon_t\right]\right\} + Var_{\varepsilon_t}\left\{E\left[\hat{C}_t \middle| \varepsilon_t\right]\right\} \\ &= \frac{C\left(E[p(\underline{v}_t, \varepsilon_t)] - E\left[p^2(\underline{v}_t, \varepsilon_t)\right]\right) + C^2 Var[p(\underline{v}_t, \varepsilon_t)]}{\left(E[p(\underline{v}_t, \varepsilon_t)]\right)^2}. \end{aligned} \tag{4.9}$$

In order to compute an estimate of the variance replace C by $\hat{C}$ as in (4.6) or (4.7).

Analytical Model for Detection Probability, $p(\underline{v}, \varepsilon)$

A candidate, and convenient, parametric analytical model is the following; described in (4.2):

$$p(\underline{v}, \varepsilon) = \exp\left[-\varepsilon\left(-\ln p_0(\underline{v})\right)\right] \tag{4.10}$$

where we can let

$$-\ln p_0(\underline{v}) = \exp\left[\beta_0 + \sum_{i=1}^{I} \beta_i v_i\right] \equiv \exp\left[\beta_0 + \underline{\beta}\,\underline{v}\right]. \tag{4.11}$$

Note that one can explicitly evaluate all expectations in (4.9) for distributions whose *Laplace transforms* are available; this is a wide class that includes the gamma, positive stable, inverse Gaussian, and many others, including convex mixtures of the above. We give only the gamma example.

Gamma. Here

$$E\left[e^{-\varepsilon s}\right] = 1 \Big/ \left(1 + \frac{1}{\alpha} s\right)^{\alpha}, \quad \alpha > 0, \tag{4.12}$$

and $E[\varepsilon] = 1$, $Var[\varepsilon] = 1/\alpha$.

The correction factor for the mean in (4.7) is seen to be

$$E\left[p(\underline{v}, \varepsilon)\right] = 1 \Big/ \left[1 + \frac{1}{\alpha} e^{\left(\beta_0 + \sum \beta_i v_i\right)}\right]^{\alpha} \tag{4.13}$$

which, for $\alpha = 1$, is just a logistic regression function; other values of $\alpha > 0$ simply generalize the model, possibly usefully. By simple analogy terms occurring in the variance are evaluated

$$E\left[p^2(\underline{v}, \varepsilon)\right] = 1 \Big/ \left[1 + \frac{1}{\alpha} e^{\left(\beta_0 + \underline{\beta}\,\underline{v}\right)2}\right]^{\alpha}; \tag{4.14}$$

$Var\left[p(\varepsilon)\right] = E\left[p^2(\underline{v}, \varepsilon)\right] - \left(E\left[p(\underline{v}, \varepsilon)\right]\right)^2$ can be evaluated using (4.14) and (4.13).

Explaining Variability Alternatively

The previous model allowed for variability in detection probability to be represented by deterministic explanatory/regression variables, $\boldsymbol{v}$, and the initial binomial model to be extended by additionally randomizing simultaneously all conditional detection probabilities on a scan/glimpse. This device can be used to represent the effects of *simultaneous* random environmental variations unexplained by regression.

An alternative possibility is to explain variations *between* potential targets (e.g. because of different orientations or partial cover) by randomly introducing an ε-effect *for each target*; an independent and identically distributed assumption is convenient and parsimonious. This approach will be pursued in later work.

Now describe the *occupancy* of a node or arc, denoted by n, by the numbers of units of different types (e.g., heavy armor brigades, light armor brigades, etc.); $U_i(n; t)$ is the number of such units of type i ($i = 0, 1, 2, \ldots, I$) at time t at node n; $U_0(n; t)$ can designate the empty node if necessary.

Within a unit of type i, there may be several asset types, such as tanks, armored personnel carriers, etc. Assets may also refer to subunits (sunits) such as tank companies. Agree that distinguishable assets occur in J classes, and that a unit of type i has a mean number of assets of type j ($j = 1, 2, \ldots, J$) equal to $\alpha_{ij}(n,t)$, and a variance $\sigma_{ij}^2(a;n,t)$. The initial values of the mean number of assets of type j for a unit of type i may be taken from the Table of Organization and Equipment (TOE) for that type of unit. Furthermore, adopt the provisional model that the actual number of assets of type j possessed by a particular randomly selected unit is a random variable, $A(i, j, n, t)$ with distribution function $F_{ij}(x; n, t)$. When convenient, and for illustration, we take $A_k(i, j; n, t)$, $k = 1, 2, \ldots, U_i$, i.e., the numbers of assets of type j owned by the U_i copies on Node n to be independent and normally/Gaussian distributed. The time-dependent parameters α_{ij} and σ_{ij}^2 can reflect the fact that a campaign has been in progress for some time and attrition has occurred and is subject to change. Let

$$\overline{A}(i,j,n,t)=\sum_{\ell=1}^{U_i}A_\ell(i,j,n,t)$$

denote the total number of assets of type j possessed by units of type i at node n at time t.

It is convenient to refer to the vector of distributions of typical asset-type numbers for a particular unit type as the *signature* (or *asset signature*) of the unit type. Note that signatures of different individual units of the same type will inevitably differ if their various asset counts differ, as could well happen. Let

$$\overline{A}(j,n,t)=\sum_{i=1}^{I}\sum_{\ell=1}^{U_i}A_\ell(i,j,n,t)$$

the total number of assets of type j at the node at time t.

Finally, let $\boldsymbol{S}_j(s, n; t)$ denote the total count of assets of type j by sensors of type s ($s = 1, 2, \ldots, S$) at node n at time t. $\boldsymbol{S}_j$ represents the quantitative *perception* of the opponent's type j asset level.

In this section and the next two, we present and more thoroughly investigate models for sensor observations of assets present on nodes or arcs that are based on the binomial and multinomial distributions. We consider the behavior of approximate procedures for updating the estimate of the number of units or assets on the node/arc using the binomial and multinomial observation models. The procedures are based on approximating the sampling distribution of the estimate of the number of units by a normal distribution. The paper by Hall (1994) gives some insight into when these approximations may be appropriate, but their convenience for the purpose of minimizing computation time is of overriding concern in the context of a computer model.

4.1 Combining Binomial Observations

A model for two independent observations of a unit with r assets by two different sensors is as follows. Let X_1 and X_2 be independent random variables having binomial distributions with common number of trials r and success probabilities p_1 and p_2. The problem is to estimate r from observations x_1 and x_2.

4.2 The Likelihood Approach

The likelihood function is

$$L(r;x_1,x_2)=\prod_{i=1}^{2}\binom{r}{x_i}p_i^{x_i}(1-p_i)^{r-x_i}. \tag{4.15}$$

An approximate maximum likelihood estimate for r can be obtained by solving

$$\frac{L(r;x_1,x_2)}{L(r-1;x_1,x_2)}=1 \tag{4.16}$$

for r. This results in the equation

$$\frac{r}{r-x_1}(1-p_1)\frac{r}{r-x_2}(1-p_2)=1. \tag{4.17}$$

Thus, r satisfies the quadratic equation

$$r^2[1-(1-p_1)(1-p_2)]-r(x_1+x_2)+x_1x_2=0. \tag{4.18}$$

If there are 3 observations a similar argument would result in r satisfying a cubic equation, etc. We get an (approximate) maximum likelihood estimate, with the good properties of the m.l.e., but must solve an awkward equation. This may not be a problem for a small number (especially one or two), but if many must be done quickly, it may be. We look for an alternative.

4.3 A Moment Estimator and the Normal Approximation

The moment estimator for r is seen to be

$$\hat{r}_i = \frac{X_i}{p_i}. \tag{4.19}$$

Note that

$$E[\hat{r}_i] = \frac{r_i p_i}{p_i} = r_i \tag{4.20}$$

and so $\hat{r}_i$ is unbiased, as it was designed to be. Further,

$$Var[\hat{r}_i] = \frac{1}{p_i^2} Var[X_i] = \frac{r_i p_i (1-p_i)}{p_i^2} = \frac{r_i (1-p_i)}{p_i}. \tag{4.21}$$

Approximate the distribution of $\hat{r}_i$ by a normal distribution with mean $\mu(i) = \frac{x_i}{p_i}$ and variance $v(i)^2 = x_i (1-p_i)/p_i^2$; if $x_i = 0$ then set $v(i)^2 = v_0^2 > 0$ where v_0^2 is chosen by the analyst. To obtain a combined estimate of r from $\hat{r}_1$ and $\hat{r}_2$ take the weighted average with weights the inverse variances:

$$\hat{r} = \frac{\frac{\hat{r}_1}{v(1)^2} + \frac{\hat{r}_2}{v(2)^2}}{\frac{1}{v(1)^2} + \frac{1}{v(2)^2}}. \tag{4.22}$$

The estimated variance is

$$\hat{\sigma}^2 = \left[\frac{1}{v(1)^2} + \frac{1}{v(2)^2} \right]^{-1}. \tag{4.23}$$

Figures 15 – 17 present results of a simulation experiment with 499 replications. For each replication two independent binomial random numbers are drawn, each with 20 trials and one having probability of success 0.5 and the other having probability 0.9. In each replication, the estimate obtained by solving (4.18) is computed and the estimate (4.22) is computed.

Figure 15 displays a histogram of the differences of the likelihood estimate and moment estimate. Note that the likelihood estimate tends to be somewhat larger than the moment estimate. Figure 16 (respectively 17) displays the histogram of the likelihood (respectively moment) estimates. The moment estimator appears to give satisfying results, although likelihood is probably superior.

4.4 Approximate Normal Updating Using Binomial Observations

Let the current estimate of the number of subunits or assets on a node, r, have a normal distribution with mean μ_t and variance σ_t^2. Suppose a new binomial observation, X_{t+1}, arrives with parameters r trials (there are, and have been, r units present) and probability of detection p. We assume the updated estimate has a normal distribution with mean

$$\mu_{t+1} = \frac{\dfrac{\mu_t}{\sigma_t^2} + \dfrac{x_{t+1}/p}{v_{t+1}^2}}{\dfrac{1}{\sigma_t^2} + \dfrac{1}{v_{t+1}^2}} \tag{4.24}$$

and variance

$$\sigma_{t+1}^2 = \left[\frac{1}{\sigma_t^2} + \frac{1}{v_{t+1}^2}\right]^{-1} \tag{4.25}$$

where

$$v_{t+1}^2 = \begin{cases} x_{t+1}(1-p)/p^2 & \text{if } x_{t+1} > 0 \\ v_0^2 & \text{if } x_{t+1} = 0 \end{cases} \tag{4.26}$$

where v_0^2 is a constant chosen by the analyst.

Note: the above algorithm is only appropriate if the latest observation is of (about) the same number of units as the earlier ones: not much has changed. But changes will occur and must be detected so the above is only a beginning.

5. Beta-Binomial Observations and Their Combination

One generalization to the simple binomial model of Section 4 is to allow the probability of detection p to be random with beta distribution having density function

$$f(p) = \begin{cases} \dfrac{p^{\alpha-1}(1-p)^{\beta-1}}{B(\alpha,\beta)} & \text{if } 0 \le p \le 1 \\ 0 & \text{if } p < 0 \text{ or } p > 1 \end{cases} \tag{5.1}$$

where

$$B(\alpha,\beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}. \tag{5.2}$$

The beta random variable has mean $m = \dfrac{\alpha}{\alpha+\beta}$ and variance $v_b^2 = \alpha\beta \Big/ \left[(\alpha+\beta)^2\left[(\alpha+\beta)+1\right]\right]$. The purpose of this generalization is to recognize variation in the actual probability of detection, p: make $\boldsymbol{p}$ a random variable.

Let $\boldsymbol{X}$ be a random variable whose conditional distribution given p is binomial with parameters r trials and probability of success p; now additionally suppose p has a beta distribution with parameters α and β

$$P\{X = x\} = \binom{N}{x}\frac{\Gamma(\alpha+\beta)\Gamma(\alpha+x)\Gamma(N-x+\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(N+\alpha+\beta)}. \tag{5.3}$$

An estimator for r is

$$\hat{r} = \frac{X}{m} \tag{5.4}$$

$$E[\hat{r}] = E\left[E[\hat{r}|p]\right] = \frac{1}{m}E[rp] = r. \tag{5.5}$$

Thus, $\hat{r}$ is unbiased. Next, evaluate the variance:

$$\begin{aligned}
Var[\hat{r}] &= E\left[Var[\hat{r}|p]\right] + Var\left[E[\hat{r}|p]\right] \\
&= E\left[\frac{1}{m^2} rp(1-p)\right] + Var\left[\frac{1}{m} rp\right] \\
&= \frac{r}{m^2}\frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)} + \frac{r^2}{m^2}\frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \\
&= \frac{r}{m^2}\frac{\alpha\beta}{(\alpha+\beta)(\alpha+\beta+1)}\left[1 + \frac{r}{(\alpha+\beta)}\right].
\end{aligned} \tag{5.6}$$

In summary,

$$Var[\hat{r}] = \frac{r}{m^2} v_b^2\left[(\alpha+\beta)+r\right] = r\frac{\beta}{\alpha(\alpha+\beta+1)}\left[\alpha+\beta+r\right].$$

Notice that this estimate depends quadratically upon the unknown r, rather than linearly as in the plain-vanilla binomial case.

5.1 Approximate Normal Updating Using Beta-Binomial Observations

Assume the current estimate of the number of subunits on a node, r, has a normal distribution with mean μ_t and variance σ_t^2. Suppose a new beta-binomial observation, X_{t+1}, occurs.

The estimator for r using X_{t+1} is $\hat{r}_{t+1} = X_{t+1}/m$ and we approximate its variance by $\hat{v}_{t+1}^2 = \hat{r}_{t+1}\,\beta\left[\alpha+\beta+\hat{r}_{t+1}\right]/\alpha(\alpha+\beta+1)$ if $x_{t+1} > 0$ and $\hat{v}_{t+1}^2 = v_0^2$ if $x_{t+1} = 0$.

We assume the updated estimate has a normal distribution with mean

$$\mu_{t+1} = \frac{\dfrac{\mu_t}{\sigma_t^2} + \dfrac{\hat{r}_{t+1}}{v_{t+1}^2}}{\dfrac{1}{\sigma_t^2} + \dfrac{1}{v_{t+1}^2}} \tag{5.7}$$

and variance

$$\sigma_{t+1}^2 = \left[\frac{1}{\sigma_t^2} + \frac{1}{\nu_{t+1}^2}\right]^{-1}. \tag{5.8}$$

5.2 Maximum Likelihood Estimate for the Number of Trials, *r*, in which the Observations are iid Beta-binomial

Let $X_1, \ldots, X_t$ be independent beta-binomial random variables with parameters r trials and beta parameters α and β. The likelihood function for r is

$$L_I(r; x_1, \ldots, x_t) = K \frac{(r!)^t}{\prod_{i=1}^{t} (r - x_i)!} \frac{\prod_{i=1}^{t} \Gamma(r - x_i + \beta)}{\Gamma(r + \alpha + \beta)^t}. \tag{5.9}$$

The maximum likelihood estimate is that integer r which maximizes L_I. An approximate solution can be found by determining that r such that

$$\frac{L_I(r; x_1, \ldots, x_t)}{L_I(r-1, x_1, \ldots, x_t)} \approx 1.$$

For $t = 2$, for $r > \max(x_1, x_2) + 1$

$$\frac{L_I(r; x_1, \ldots, x_t)}{L_I(r-1; x_1, \ldots, x_t)} = \frac{r^2}{(r - x_1)(r - x_2)} \frac{(r - x_1 + \beta - 1)(r - x_2 + \beta - 1)}{(r - 1 + \alpha + \beta)^2}. \tag{5.10}$$

Setting the ratio equal to 1 results in a cubic equation for r; the minimum solution larger than $\max(X_i)$ is selected as the reasonable point estimate.

Figures 18 and 19 display results of a simulation experiment with 99 replications. Each replication consists of 2 independent beta-binomial random numbers; the binomial number of trials is 20 and the parameters of the beta distribution are $\alpha = 5$ and $\beta = 2$.

Figure 18 (respectively Figure 19) displays a histogram of the moment estimates (respectively the maximum likelihood estimates). Note that the moment estimate is much more variable than the maximum likelihood estimate.

5.3 Common but Random Probability of Detection

Assume $X_1, \ldots, X_t$ are conditionally independent binomial random variables given p with r trials and probability of success p but p has a beta distribution with parameters α and β. Randomization of p common to all sensors could represent a common effect on them all, e.g. by weather, or cloud cover, or perhaps terrain variation. Then

$$P\{X_1 = x_1, X_2 = x_2, \ldots, X_t = x_t\}$$

$$= \left[\prod_{j=1}^{t} \binom{r}{x_j}\right] \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma\left(\alpha + \sum_{i=1}^{t} x_i\right)\Gamma\left(\beta + \sum_{i=1}^{t}(r - x_i)\right)}{\Gamma(\alpha+\beta+tr)}. \tag{5.11}$$

Thus, the likelihood function for r is of the form (K an arbitrary constant)

$$L(r; x_1, \ldots, x_t) = K \prod_{j=1}^{t} \binom{r}{x_j} \frac{\Gamma\left(\beta + \sum_i (r - x_i)\right)}{\Gamma(\alpha+\beta+tr)}. \tag{5.12}$$

For $r \geq \max_i(x_i) + 1$ and $t = 2$

$$\frac{L(r; x_1, \ldots, x_{t+1})}{L(r-1; x_1, \ldots, x_t)}$$
$$= \frac{r^2}{(r-x_1)(r-x_2)} \frac{(r - x_1 + r - x_2 + \beta - 1)(r - x_1 + r - x_2 + \beta - 2)}{(2r+\alpha+\beta-1)(2r+\alpha+\beta-2)}. \tag{5.13}$$

Setting the ratio equal to 1 results in a quartic equation for r which can be solved (numerically); the minimum solution greater than $\max(x_i)$ is the reasonable one. Numerical stability concerns suggest that the logarithm of the likelihood function be computed and Stirling's formula (cf. Feller [1968]) be used to approximate the gamma functions; that is, compute

$$\begin{aligned}\ell(r;x_1,\ldots,x_t) &= \sum_{j=1}^{t}\ln\binom{r}{x_j} + \left[\beta + \sum_i (r-x_i) - \frac{1}{2}\right]\ln\left[\beta + \sum_i (r-x_i) - 1\right] \\ &-\left[\beta + \sum_i (r-x_i) - 1\right] \\ &-\left[\alpha+\beta+tr-\frac{1}{2}\right]\ln[\alpha+\beta+tr] + [\alpha+\beta+tr-1] \\ &= \sum_{j=1}^{t}\ln\binom{r}{x_j} + \left[\beta + \sum_i (r-x_i) - \frac{1}{2}\right]\ln\left[\beta + \sum_i (r-x_i) - 1\right] \\ &-\left[\alpha+\beta+tr-\frac{1}{2}\right]\ln[\alpha+\beta+tr] + K\end{aligned}$$

where K does not depend on r.

Figures 20 – 23 report the results of simulation experiments. For each replication one random number, p, is drawn from a beta distribution with parameters α and β. For the experiments reported in Figures 20 – 21 (respectively Figures 22 – 23) two (respectively 10) independent random numbers are then drawn from a binomial distribution with 20 trials and common probability of success p. Conclusion: for cases considered, the moment estimator appears to be an adequate approximation to the maximum likelihood estimate. Note that, as mentioned earlier, the above Beta binomial model represents extra variability common to *sets* of binomial probabilities (of success). The present Beta binomial does *not* represent any extra binomial randomness *between* the outcomes of successive looks at the same unit by the same sensor (within variability).

6. The Many-Unit Type (Multinomial) Misclassification Problem

We next present a more realistic version of the models considered earlier in this report. In an actual operating environment there is presumed to be a *mélange* of different observables: members of several classes of different subunits, such as tank companies: r_i subunits of type i, $i = 1, 2, \ldots I$ in a specified part of the theater (on a given arc or at a given node). Assume that each subunit type has a probability of being detected on a sensor pass or observational opportunity; let this be d_i for subunits of type i. Furthermore, once detected the subunit is assumed correctly classified with probability c_{ii}, but incorrectly so, and to one of the recognized classes, with probability c_{ij}, $j \neq i$. It is optimistic, but convenient, to assume that $c_{ii} \approx 1$ and c_{ij}, $j \neq i$, is close to zero; this will not always be necessary but can be a useful start at times. An overly simplistic way of handling the misclassification problem is to simply assume that the probability of misclassifying i to $j \neq i$ is a small constant.

Note that each of the above parameters is likely to be specific to sensor type: $d_i(s)$, $\xi_{ij}(s)$, letting the sensors be classified as of type $s = 1, 2, \ldots, S$. And, during a particular (e.g. 2-hour) reference time period the number of opportunities that a given sensor has for "seeing" a unit of type i at a given location will vary depending upon the action(s) of the subunit(s) and the number of passes the sensor makes over the particular region under examination. This will be a programmed quantity and will be known. The actual detectability of a subunit, i.e. as influenced by the subunit present where looking occurs is unknown and must be estimated.

6.1 Model

Begin by considering one pass by a sensor over a region (e.g. arc). First, carry out *detections*: let $\boldsymbol{D}_i$ be a random variable denoting the number of detections of (sub)units of type i; if desired specify $\boldsymbol{D}_i(s)$. Assuming sampling with replacement is adequate,

$$\boldsymbol{D}_i \text{ is Binomial } (r_i, d_i). \tag{6.1}$$

We assume that $\boldsymbol{D}_i$ is not directly observable. Of course there is identity information inherent in what different sensor types see: if sensor type 1 can only see subunit type 1, sensor type 2 can only see subunit type 2, then if $d_1 = 10$ and $d_2 = 0$ it is clear what *types* of subunits are present.

Next, *classify* all of the detections according to the probabilities c_{ij}. Consequently each of the $\boldsymbol{D}_i$ (anonymous) observations is independently identified with some class j, where $j = 1, 2, \ldots, I$. Thus r_i, the unknown number of type i subunits gives rise to a multinomial distribution of numbers of values $\boldsymbol{X}_{ij}$

$$P\{\boldsymbol{X}_{i1} = x_{i1}, \boldsymbol{X}_{i2} = x_{i2}, \ldots, \boldsymbol{X}_{iI} = x_{iI}\} = \binom{r_i}{x_{i1}, \ldots x_{iI}} \prod_{j=1}^{I} p_{ij}^{x_{ij}} (1-d_i)^{r_i - \sum_{j=1}^{I} x_{ij}} \tag{6.2}$$

where $p_{ij} = d_i c_{ij}$.

But what can be observed is, say,

$$x_j = \sum_{i=1}^{I} x_{ij} \tag{6.3}$$

all observations *classified* as type j. The objective is to turn the above into a reasonable estimator of r_i, assuming that d_i and c_{ij} are known, at least initially. As stated before, there is useful classification information in $x_j(s)$, $s = 1, 2, \ldots, S$ if sensors are differentially sensitive.

Moment Estimator

Apparently a relatively quick way of estimating the parameters $r_1, r_2, \ldots, r_I$ is to use the method of moments. This may not be as efficient as other ways, but conveniently makes low computational demands.

Let $\boldsymbol{X}_{ij}(s)$ denote the number of times a sensor of type s sees a unit of type i and classifies and reports it as of type j. Then

$$E\left[\boldsymbol{X}_{ij}(s)\right] = r_i d_i(s) c_{ij}(s) \equiv r_i p_{ij}(s). \tag{6.4}$$

Now $\boldsymbol{X}_j(s) = \sum_{i=1}^{I} \boldsymbol{X}_{ij}(s)$ models the actual observed data $x_j(s)$, so the method of moments puts

$$x_j(s) = \sum_{i=1}^{I} r_i p_{ij}(s) \qquad j = 1,2,\ldots,I. \tag{6.5}$$

In matrix terms

$$\mathbf{x}(s) = \mathbf{rp}(s) \tag{6.6}$$

so, if the inverse exists

$$\hat{\mathbf{r}} = \mathbf{x}(s)\mathbf{p}^{-1}(s) \tag{6.7}$$

This is directly analogous to the previous scalar setup. But a solution of I simultaneous equations is now required. Notice that $\sum_{j=1}^{I} p_{ij}(s) = d_i(s) \leq 1$.

Approximations

The solution offered above may be statistically somewhat inefficient (use of moments, rather than likelihood or Bayes) and is also computationally troublesome. Here are some approximations; these are tentative and will be checked by simulation. They are based on the *assumption* that $p_{ij}(s)$, $j \neq i$, can often be expected to be much smaller than $p_{ii}(s)$. If not, then we are sometimes led to say, skeptically, "I have just seen a goat, but I know that the other side always dresses sheep like goats, so I will treat this as a sheep." Such is related to *decoy and deception* issues, but also to imperfect Battle Damage Assessment (BDA), and to occurrence of false targets in general.

Assuming the above to be true we can plausibly estimate r_i initially by the *first* (naive) *estimate*:

$$\hat{r}_i^{(1)}(s) = \frac{x_i(s)}{p_{ii}(s)} \tag{6.8}$$

The model for this is

$$\hat{\boldsymbol{r}}_i^{(1)}(s) = \frac{\boldsymbol{X}_i(s)}{p_{ii}(s)}. \tag{6.9}$$

Now take expectations:

$$\begin{aligned}
E\left[\hat{r}_i^{(1)}(s)\right] &= \frac{1}{p_{ii}(s)} E\left[\sum_{k=1}^{I} X_{ki}(s)\right] = \frac{1}{p_{ii}(s)} \sum_{k=1}^{I} r_k p_{ki}(s) \\
&= \frac{r_i p_{ii}(s)}{p_{ii}(s)} + \sum_{k \neq i} r_k \frac{p_{ki}(s)}{p_{ii}(s)} \qquad (6.10) \\
&= r_i + \sum_{k \neq i} r_k \frac{p_{ki}(s)}{p_{ii}(s)}.
\end{aligned}$$

Now since $p_{ki} << p_{ii}$ the second term can well be small (unless $r_k >> r_i$). So a possible way to correct (and improve) the initial estimator is to approximately remove the estimated bias, i.e. compute the *second estimate*:

$$\hat{r}_i^{(2)}(s) = \max\left(0, \hat{r}_i^{(1)}(s) - \sum_{k \neq i} \hat{r}_k^{(1)}(s) \frac{p_{ki}(s)}{p_{ii}(s)}\right). \qquad (6.11)$$

It might be worth iterating this to obtain $\hat{r}_i^{(3)}(s)$, etc., but this step will not be taken. To obtain an estimate for the variance of $\hat{r}_i^{(2)}(s)$, note that

$$Var\left[X_j(s)\right] = \sum_{i=1}^{I} r_i p_{ij}(s)\left[1 - p_{ij}(s)\right]$$

which can be estimated as

$$\hat{Var}\left[X_j(s)\right] \cong \sum_{i=1}^{I} \hat{r}_i^{(2)}(s) p_{ij}(s)\left[1 - p_{ij}(s)\right] \qquad (6.12)$$

for example.

Approximately,

$$\begin{aligned}
Var\left[r_i^{(2)}(s)\right] &\approx Var\left[r_i^{(1)}(s)\right] + \sum_{k \neq 1} Var\left[r_k^{(1)}(s)\right]\left[\frac{p_{ki}(s)}{p_{ii}(s)}\right]^2 \\
&\approx \frac{\hat{Var}\left[X_i(s)\right]}{p_{ii}^2(s)} + \sum_{k \neq i} \frac{\hat{Var}\left[X_k(s)\right]}{p_{kk}^2(s)}\left[\frac{p_{ki}(s)}{p_{ii}(s)}\right]^2;
\end{aligned} \qquad (6.13)$$

if $X_j(s) = 0$ set $Var[X_j(s)] = v_0$ where v_0 is chosen by the analyst. Of course this last is *ad hoc*; its properties are best understood by use of some simulation.

6.2 Approximate Normal Updating Using Multinomial Observations

Let the current estimate of the number of subunits of type $i = 1, \ldots, I$ on a node, r_i, have independent normal distributions with mean $\mu_t(i)$ and variance $\sigma_t^2(i)$. Suppose a new multinomial observation X_{t+1} arrives from a situation with parameters r_i, $p_{ij}(s)$, for $i = 1, \ldots, I$ and $j = 1, \ldots, I$. As usual r_i is unknown, but $p_{ij}(s)$ is assumed known.

We assume the updated estimators have independent normal distributions with means

$$\mu_{t+1}(i) = \frac{\dfrac{\mu_t(i)}{\sigma_t^2(i)} + \dfrac{\hat{r}_i}{v_{t+1}^2(i)}}{\dfrac{1}{\sigma_t^2(i)} + \dfrac{1}{v_{t+1}^2(i)}} \tag{6.14}$$

and variance

$$\sigma_{t+1}^2(i) = \frac{1}{\dfrac{1}{\sigma_t^2(i)} + \dfrac{1}{v_{t+1}^2(i)}} \tag{6.15}$$

where $\hat{r}_i$ is an estimate of r_i and $v_{t+1}^2(i)$ is an estimate of the variance of $\hat{r}_i$. The estimate of the variance uses (6.12) – (6.13).

A Simulation Experiment

In this section results of a simulation experiment are reported.

In each replication of the simulation two multinomial random numbers are generated. Each has number of trials (r_1, r_2), probabilities of detection (d_1, d_2) and probabilities of classification $(c_{11}, c_{12}, c_{21}, c_{22})$. The observed data are

$$X_1(1;s) = X_{11}(1;s) + X_{21}(1;s); X_1(2;s) = X_{11}(2;s) + X_{21}(2;s)$$

and

$$X_2(1;s) = X_{12}(1;s) + X_{22}(1;s); X_2(2;s) = X_{12}(2;s) + X_{22}(2;s).$$

Two procedures for combining the two observations are considered. In one, called *Average*, the average of the two observations is computed and the estimate (6.11) is computed using the average $\overline{X}_1, \overline{X}_2$ where

$$\overline{X}_1 = [X_1(1;s) + X_1(2;s)]/2.$$

Figures 24 and 25 present histograms of the moment estimate using one observation ($X_1(1; s)$, $X_2(1; s)$) obtained by solving simultaneous equations (6.6) and histograms of the approximate moment estimate (6.11). The two estimates appear to be reasonably close with the approximate moment estimate being slightly lower. Table 6.1 presents statistics for the simulation results.

Table 6.1
1 observation
$d_1 = 0.7, d_2 = 0.9, c_{11} = 0.8, c_{12} = 0.2, c_{21} = 0.1, c_{22} = 0.9$
500 replications: $r_1 = 10, r_2 = 15$

Proc	Mean of Estimates		Variance of Estimates	
	$\hat{r}_1$	$\hat{r}_2$	$\hat{r}_1$	$\hat{r}_2$
Simul Eqtn	10.3	14.8	14.8	7.34
Approx	10.0	14.4	14.0	6.9

Figures 26 – 27 present histograms of the estimates resulting from the approximate normal procedure for combining two multinomial observations ($X_1(1; s)$, $X_2(1; s)$), ($X_1(2; s)$, $X_2(2; s)$); the estimates of r_i combined are those resulting from (6.11), (6.12), (6.13), (6.14), (6.15). The figures also present histograms of first averaging the two multinomial observations and then applying estimator (6.11) – (6.13). The two procedures appear to yield comparable histograms. The estimated variances are computed using (6.13). The mean of the estimated variance is always smaller than the variance of the estimates. Table 6.2 presents statistics of the simulation results.

Table 6.2
2 observations
$d_1 = 0.7, d_2 = 0.9, c_{11} = 0.8, c_{12} = 0.2, c_{21} = 0.1, c_{22} = 0.9$
500 replications: $r_1 = 10, r_2 = 15$

Proc	Mean of Estimates		Variance of Estimates		Mean of Variance Estimates	
	$\hat{r}_1$	$\hat{r}_2$	$\hat{r}_1$	$\hat{r}_2$	$\hat{r}_1$	$\hat{r}_2$
Ave	9.8	14.7	6.1	3.6	5.8	2.8
Normal	9.4	14.9	6.7	6.0	5.7	4.9

Table 6.3 records statistics from a simulation with 500 replications. For each estimate in a replication the following confidence interval is computed: $\left[\hat{r}_i - 2\sqrt{\hat{Var}\,\hat{r}_i},\ \hat{r}_i + 2\sqrt{\hat{Var}\,\hat{r}_i}\right]$; this confidence interval should have coverage at least 95%. Table 6.3 records the fraction of confidence intervals which cover the true value of r_i. Note that the fraction tends to be a little smaller than 0.95.

Table 6.4 reports results of a similar simulation but one in which $r_1 = 5$ and $r_2 = 8$. Note the fraction of confidence intervals that cover the true r_i tend to be smaller than 0.95. This behavior is due to the greater likelihood that the estimate of r_i is 0.

Table 6.3
Statistics of Simulation
$d_1 = 0.7, d_2 = 0.9, c_{11} = 0.8, c_{22} = 0.9, c_{12} = 0.2, c_{21} = 0.1$
500 replications
$r_1 = 10, r_2 = 15$

No. of Obs	Type of Unit	Estimator	Mean of Est (Std Dev)	Mean of Est Std Dev	Fraction of intervals $\left[\hat{r}_i \pm 2\sqrt{\hat{Var}\,\hat{r}_i}\right]$ that covers true value
1	1	Approximate Moment	9.8 (3.5)	(3.4)	0.93
	2		14.8 (2.7)	(6.6)	1.0
2	1	Normal Updating with Approx Moment est.	9.6 (2.5)	(2.4)	0.93
	2		14.4 (2.3)	(2.2)	0.94
2	1	Average obs. first then approx Moment Est	9.9 (2.4)	(2.4)	0.96
	2		14.5 (1.9)	(1.6)	0.91

Table 6.4
Simulation Statistics
$r_1 = 5, r_2 = 8, d_1 = 0.7, d_2 = 0.9, c_{11} = 0.8, c_{22} = 0.9$
500 replications

No. of Obs	Type of Unit	Estimator	Mean of Est (Std Dev of Est)	Mean of Est Std Dev	Fraction of $\left[\hat{r}_i - 2\sqrt{\hat{Var}\,\hat{r}_i},\ \hat{r}_i + 2\sqrt{\hat{Var}\,\hat{r}_i}\right]$ that covers true r_i
1	1	Approximate Moment	5.0		0.93
			(2.6)	(2.5)	
	2		7.8		1
			(1.9)	(4.7)	
2	1	Normal Updating of Approx Moment est.	4.6		0.87
			(1.9)	(1.7)	
	2		7.8		0.94
			(1.7)	(1.6)	
2	1	First average obs. then use Approx Mom. Est.	4.9		0.91
			(1.9)	(1.7)	
	2		7.8		0.91
			(1.4)	(1.2)	

Observe that the coverage fraction in the above table is nearly always too low, i.e. is not conservative. Practically, one might simply change from a multiplier by 2 to something a bit larger, e.g. from the Student's *t* distribution.

All of the above could be redone in a Beta-mixed context.

7. Some Approaches to Updating Perception At a Node Using Information from Neighboring Nodes

7.1 Introduction

Describe the occupancy of a node or arc by the number of units of different types that occupy the node or arc; the node or arc may well be empty. Since units move along arcs and nodes, the occupancy of a node/arc changes over time. Each unit may have several types of assets, e.g. tanks, soldiers, APC's (Armored Personnel Carriers), etc. and different types of subunits, e.g. tank companies. A protagonist gains information about the node occupancies through the use of sensors which give him error-prone observations. In an idealistic and theoretical sense it would be desirable for a protagonist in a theater-level campaign to possess the joint distribution of the types and numbers of units occupying all relevant nodes, along with their assets, at every node for all times t. This distribution will be influenced by the protagonist's allocation of sensor assets, by his own force maneuver as well as by the opponent's maneuvers to accomplish a planned course of action.

However, the development of an automated systemic model of a theater campaign must realistically stop short of a completely integrated joint probabilistic assessment of node occupancy; computational burdens are simply too great.

Algorithms for updating perception of the numbers of assets/subunits and numbers and types of units at a node, using local information for that node, have been discussed previously; see Appendix A and Sections 4 – 6. In this section we investigate the behavior of alternative perception updating algorithms. Some of these algorithms use information from neighboring nodes or arcs.

7.2 The Model

For simplicity of discussion and notation we will assume for the present that there is one type of unit, e.g. brigade, with only one asset type. Let $K(n, t)$ be the number of units that occupy node n at time t and let $A(n, t)$ be the number of assets or subunits at node n at time t. Assume the prior distribution for the number of assets given the number of units with no sensor observations is

$$P\{A(n,0) \in da | K(n,0) = k\} = \frac{1}{\sqrt{2\pi}\sqrt{k}\sigma_a} \exp\left\{-\frac{1}{2}\frac{(a-k\gamma)^2}{k\sigma_a^2}\right\} \qquad (7.1)$$

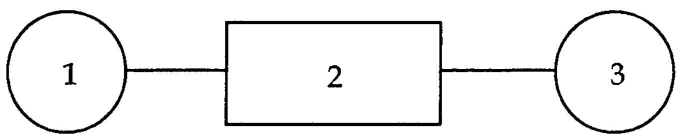
where γ is the mean number of assets associated with one unit and σ_a^2 is the variance of the number of assets associated with one unit; γ is derivable from the Table of

Organization and Equipment (TOE) for the unit plus any information available concerning unit campaign experience.

We assume one kind of sensor which counts assets with

$$P\left\{S \in dx \middle| A(n)=a\right\} = \frac{1}{\sqrt{2\pi\tau}} \exp\left\{-\frac{1}{2}\left(\frac{x-a}{\tau}\right)^2\right\}. \tag{7.2}$$

For illustration purposes suppose there are three nodes.



Suppose it is known that units at node 1 will, after a random time T_1, move to node 2 and after a further random time T_2, move to node 3.

The problem is to use sensor observations to *estimate the number of units* and number of assets at each node. In general, there will not be enough sensors to observe all three nodes at all times.

Let the marginal distribution of the number of units on node n at t, given all sensor observations up to t be

$$\pi(k;n,t) = P\left\{K(n,t)=k \middle| \text{all sensor observations during } [0,t]\right\} \tag{7.3}$$

and let the marginal distribution of the number of assets at node n, given sensor observations and units be

$$\begin{aligned} dF_k(a;n,t) &= P\left\{A(n,t) \in da \middle| \text{all sensor observations during } [0,t], K(n,t)=k\right\} \\ &\equiv \frac{1}{\sqrt{2\pi v_k(n,t)}} \exp\left\{-\frac{1}{2}\left(a-m_k(n,t)\right)^2 / v_k^2(n,t)\right\}. \end{aligned} \tag{7.4}$$

In the remainder of the paper we present algorithms for updating π and F_k as more information becomes available.

7.3 Updating Using Same Node Information

Suppose there is a sensor observation $x(n, t+1)$ of the assets at node n at time $t+1$. One can update the perception at node n using only the sensor observation at node n as follows.

$$v_k(n,t+1) = \left[\frac{1}{v_k^2(n,t)/\alpha(a)} + \frac{1}{\tau^2}\right]^{-1} \tag{7.5}$$

$$m_k(n,t+1) = \frac{\dfrac{m_k(n,t)}{v_k^2(n,t)/\alpha(a)} + \dfrac{x(n,t+1)}{\tau^2}}{\dfrac{1}{v_k^2(n,t)/\alpha(a)} + \dfrac{1}{\tau^2}} \tag{7.6}$$

$$\begin{aligned} \pi(k,n;t+1) &= C\pi(k,n,t)\frac{1}{\sqrt{2\pi}\sqrt{\tau^2 + \dfrac{v_k^2(n,t)}{\alpha(a)}}} \exp\left\{-\frac{1}{2}\left[\frac{\left(x(n,t+1) - m_k(n,t)\right)^2}{\tau^2 + \dfrac{v_k^2(n,t)}{\alpha(a)}}\right]\right\} \\ &\equiv C\pi(k,n,t)w(k;n,t,x) \end{aligned} \tag{7.7}$$

where $0 < \alpha(a) \le 1$ is a constant discount value. If $\alpha(a) = 1$, then the procedure is the same as that in Appendix A and Sections 4 – 6.

If there is no sensor observation of node n at time $t+1$, then set

$$m_k(n,t+1) = m_k(n,t) \tag{7.8}$$

$$v_k^2(n,t+1) = v_k^2(n,t)/\alpha_0(a) \tag{7.9}$$

for a discount constant $0 < \alpha_0(a) \le 1$.

7.4 Updating Using Neighboring Node Perception

Since it is assumed from the combat situation of the section that the units that are on node 1 will eventually move to node 2, it would be advantageous to incorporate this information in the updating.

a. Modifying the Prior of the Distribution of Numbers of Units

One consequence of a Bayesian procedure is that as more evidence accumulates for a particular number of units at a node, the procedure becomes more sluggish in responding to a changing situation. One way to overcome this sluggishness is to modify the prior of the number of units that are at a particular node. Specifically, suppose there is a sensor observation $x(n, t+1)$ at node n at time $t+1$. Before applying the updating procedure of (7.5) – (7.7) modify $\pi(k, n, t)$ as follows. For $n > 1$

$$\tilde{\pi}(k;n,t) = (1-\alpha(p))\,\pi(k;n,t) + \alpha(p)\,\pi(k;n-1,t) \tag{7.10}$$

where $0 < \alpha(p) < 1$ is a constant. The constant $\alpha(p)$ may or may not be the same as the constant discount factor $\alpha(a)$. If $n = 1$ (node 1), then

$$\tilde{\pi}(k;n,t) = (1-\alpha(p))\,\pi(k;n,t) + \alpha(p)\,\varepsilon_0(k)$$

with

$$\varepsilon_0(k) = \begin{cases} 1 & \text{if } k = 0 \\ 0 & \text{if } k > 0. \end{cases} \tag{7.11}$$

Then, as before

$$\pi(k;n,t+1) = C\tilde{\pi}(k;n,t)w(k;n,t,x) \tag{7.12}$$

where $w(k; n, t, x)$ is defined in (7.7).

If there is no sensor observation, one possibility for updating is

$$\pi(k;n,t+1) = (1-\alpha_0(p))\,\pi(k;n,t) + \alpha_0(p)u(k) \tag{7.13}$$

where u is a discrete uniform distribution over the possible number of units and $0 \le \alpha_0(p) < 1$ is a constant.

b. Modifying the Distribution of the Number of Assets

The procedure of Section 7.3 updates the conditional mean, m_k, and variance, v_k^2, of the number of assets at the node given the number of units at the node is k for all possibilities of number of units. It may be advantageous to occasionally reset the moments of the number of assets to the TOE (Table of Organization and Equipment) values, possibly reduced by perceived attrition. One procedure to reset the moments is described in Appendix A.

Another procedure follows. Associate with each node, not only the posterior mean and variance at time t for k units, $m_k(t)$ and $v_k^2(t)$ but also the TOE mean and variance $k\gamma$ and $k\sigma_a^2$ of the number of assets at the node given the number of units that are at the node is k. Let

$$J(k,n,t)=\begin{cases}1 & \text{if the sensor observation at time } t \text{ belongs to the prior distribution of the number of assets for } k \text{ units at node } n \text{ at time } t\\ 0 & \text{if the sensor observation at time } t \text{ belongs to the TOE distribution of the number of assets for } k \text{ units at node } n \text{ at time } t.\end{cases}$$

Let

$$\pi_J(j,k;n,t)=P\{J=j,K(n,t)=k \mid \text{all sensor observations during } [0,t]\}.$$

Suppose there is a sensor observation $x(n, t+1) = x$ of the assets at node n at time $t+1$, then a similar procedure to (7.7) can be used to update π_J; i.e.

$$\pi_J(j,k;n,t+1)=C\pi_J(j,k;n,t)w_J(j,k;n,t,x) \tag{7.14}$$

with

$$w_J(j,k;n,t,x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\sqrt{\tau^2+\frac{v_k^2(n,t)}{\alpha(a)}}}\exp\left\{-\dfrac{1}{2}\left[\dfrac{\left(x-m_k(n,t)\right)^2}{\tau^2+\frac{v_k^2(n,t)}{\alpha(a)}}\right]\right\} & \text{if } j=1\\[2ex] \dfrac{1}{\sqrt{2\pi}\sqrt{\tau^2+k\sigma_a^2}}\exp\left\{-\dfrac{1}{2}\left[\dfrac{(x-k\gamma)^2}{\left(\tau^2+k\sigma_a^2\right)}\right]\right\} & \text{if } j=0.\end{cases} \tag{7.15}$$

Further,

$$P\{K=k|\text{all observations in }(0,t+1]\}=\sum_j \pi_J(j,k;n,t+1) \tag{7.16}$$

$\Bigg($ respectively,

$$P\{J=j|\text{all observations in }(0,t+1]\}=\sum_k \pi_J(j,k;n,t+1)\Bigg) \tag{7.17}$$

gives the posterior probability of there being $K=k$ units at node n at time $t+1$ (respectively the posterior probability that the TOE moments summarize the observation better than the prior moments).

Finally, if $P\{J=0|\text{all observations in }(0,t+1]\}$ is larger than some constant α_J (e.g. $\alpha_J = 0.9$) one can reset the posterior moments equal to the TOE values; that is

$$m_K(n,\, t+1) = k\gamma$$

and

$$v_k^2(n,t+1)=k\sigma_a^2 .$$

In addition, one can modify the prior distribution of the number of units *a la* (7.10) – (7.13). For example, for $n > 1$

$$\tilde{\pi}_J(j,k;n,t)=(1-\alpha(p))\pi_J(j,k;n,t)+\alpha(p)\pi_J(j,k;n-1,t)$$

where $0 < \alpha(p) < 1$.

7.5 A Minimal Updating Procedure

A minimal updating procedure might always use the TOE values for the mean and variance of the number of assets and treat the problem of inference of the number of units at a node as a classification problem.

In particular, let $\pi_M(k, n; t)$ be the estimate of the probability that there are k units at node n at time t given all the sensor observations up through time t. Suppose there is a sensor observation $x(n, t+1)$ of the assets at node n at time $t+1$. Modify the prior distribution of the number of units at the node as in (7.10) – (7.11); i.e.

$$\tilde{\pi}_M(k,n;t) = (1-\alpha(p))\tilde{\pi}_M(k,n;t) + \alpha(p)\tilde{\pi}_M(k,n-1;t) \tag{7.18}$$

for $n > 1$ and

$$\tilde{\pi}_M(k,1;t) = (1-\alpha(p))\pi_M(k,1;t) + \alpha(p)\varepsilon_0(k). \tag{7.19}$$

To compute the estimate of the distribution of the number of units at the node

$$\pi_M(k,n;t+1) = C\tilde{\pi}_M(k,n;t+1)\frac{1}{\sqrt{2\pi}\sqrt{\tau^2 + k\sigma(a)^2}}\ \exp\left\{-\frac{1}{2}\frac{(x(n,t+1)-k\gamma)^2}{\tau^2 + k\sigma(a)^2}\right\} \tag{7.20}$$

where C is a normalization constant.

7.6 Examples

The example of Section 7.2 is simulated. The random times, T_i, $i = 1, 2$ of occupancy at node i are independent having normal distributions with mean 5 and standard deviation 1. The standard deviation of the sensor is τ. The TOE mean for one unit $\gamma = 100$ and the standard deviation $\sigma_a = 10$. There is one sensor measurement at node 1 at times 1–7. There is one sensor measurement at node 2 at times 5–12. There is one measurement at node 3 at times 10–20. The simulation replication is for a total of 20 time units. The true number of units is 2. If there are no observations of a node during a reporting cycle, the distribution of the number of units at the node is a discrete uniform over 0, 1, 2, 3.

The following statistics are collected for each replication. If a sensor observation occurs at node n, the posterior probability of the true number of units at the node is collected. Also collected is the posterior mean number of assets, e.g. $\bar{a} = \sum m_k(n,t)\pi(k;n,t)$. Finally, the average posterior probability of the correct unit is computed over all observations of all nodes over all observation times is computed. Also computed is the average absolute deviation between the posterior mean number of assets and the true number of assets at the node over all nodes with observations and observation times.

There are 20 replications for each simulation. Table 7.1 displays means and standard deviations of the replication statistics over the 20 replications. The results indicate that modifying the prior of the number of units as in (7.10) is associated with the greatest improvement in the updating procedure for the perception of the number of units at a node. If the sensor observation is very good, with small standard error, there is the suggestion that resetting the conditional posterior moments of the number of assets to their TOE values is worthwhile. Finally, the minimal updating procedure with the moments of the distribution of assets always equal to the TOE values appears to be adequate.

7.7 More Than One Observation Per Reporting Cycle

In the previous subsections of this section there has been 1 observation per reporting cycle. In this subsection suggestions for updating procedures with more than one observation per reporting cycle are given.

Since it is known that the ground truth at a node will change, it may not be prudent to weight all observations obtained during a cycle equally. Presumably older observations should be down-weighted.

Suppose that observations $x_1, x_2, \ldots, x_O$ are taken during a reporting cycle at times t_1, $t_2, \ldots, t_O$ with $t_1 \leq t_2 \leq \ldots \leq t_O$. Observation x_i is an observation from a sensor with standard deviation τ_i. The reporting cycle is assumed to start at time 0 and end at time t.

a. Summarizing the Sensor Observations

One possibility for updating is to summarize the O observations and then use the TOE classification procedure of Section 7.5 to update the estimate of the probability that there are k units at the node at the end of the reporting cycle.

Table 7.1
Summary Statistics for Updating Procedures

τ	Smoothing Parameter for Number of Units $\alpha(p)$ (7.10)	Smoothing Parameter for Number of Assets $\alpha(a)$ (7.5)	Posterior Probability Used	Average of Mean Posterior Probability of Correct Number of Units (std dev. over simulation repl.)	Average of Mean Deviation of Posterior Mean Number of Assets from Correct Number (std dev. over simulation repl.)
20	0	1	π	0.71 (0.08)	41.3 (5.9)
20	0	1	π_J	0.72 (0.06)	40.9 (8.3)
20	0	1	π_M	0.76 (0.07)	44.3 (9.8)
20	0.1	1	π	0.98 (0.03)	25.7 (3.8)
20	0.1	1	π_J	0.98 (0.02)	7.37 (3.4)
20	0.1	1	π_M	0.99 (0.03)	12.1 (7.4)
50	0.1	1	π	0.82 (0.06)	28.6 (6.3)
50	0.1	1	π_J	0.82 (0.08)	25.3 (9.3)
50	0.1	1	π_M	0.81 (0.11)	22.5 (8.7)
50	0.1	0.9	π	0.79 (0.08)	28.8 (7.0)
50	0.1	0.9	π_J	0.79 (0.14)	25.7 (8.8)
50	0	0.9	π	0.61 (0.08)	49.0 (6.6)
50	0	0.9	π_J	0.57 (0.12)	54.0 (8.7)

One weighted average estimate of the number of assets at the node at the end of the update cycle is

$$\hat{a} = \frac{\displaystyle\sum_{i=1}^{O} \frac{x_i}{\tau_i^2\, \alpha(a)^{-(t-t_i)}}}{\displaystyle\sum_{i=1}^{O} \frac{1}{\tau_i^2\, \alpha(a)^{-(t-t_i)}}}. \tag{7.21}$$

This procedure will down-weight older observations. The parameter $0 < \alpha(a) \leq 1$ is a smoothing parameter. The smaller $\alpha(a)$ is the less the older observations will influence the estimated number of assets, $\hat{a}$.

$$\mathrm{Var}\left[\hat{A}\right] = \frac{\displaystyle\sum_{i=1}^{O} \frac{\tau_i^2}{\left(\tau_i^2\, \alpha(a)^{-(t-t_i)}\right)^2}}{\left(\displaystyle\sum_{i=1}^{O} \frac{1}{\tau_i^2\, \alpha(a)^{-(t-t_i)}}\right)^2}. \tag{7.22}$$

Note that if $\alpha(a) = 1$ and $\tau_i = \tau$, then $\hat{a}$ is the sample average of the observations and $\mathrm{Var}\left[\hat{A}\right] = \frac{\tau^2}{O}$.

To update the probability of k units at the node at the end of the reporting cycle

$$\pi_O(k;n,t) = C\tilde{\pi}_O(k;n,t-\Delta)\frac{1}{\sqrt{2\pi}\sqrt{k\sigma(a)^2 + \mathrm{Var}\left[\hat{A}\right]}}\exp\left\{-\frac{1}{2}\frac{(\hat{a}-k\gamma)^2}{k\sigma(a)^2 + \mathrm{Var}\left[\hat{A}\right]}\right\} \tag{7.23}$$

where C is a normalization constant, $\tilde{\pi}_O$ is the (possibly modified) probability from the previous reporting cycle, and Δ is the length of the reporting cycle.

It may be advantageous to have the value of the smoothing parameter $\alpha(a)$ depend on evidence that the observations, $x_1, \ldots, x_O$ are not from the same distribution. If $x_1, \ldots, x_O$ come from a distribution with the same mean, then

$$D_{i+1} = X_{i+1} - X_i \tag{7.24}$$

has mean 0 and variance $\tau_{i+1}^2 + \tau_i^2$.

Let $1 \geq \alpha_N(a) > \alpha_C(a) > 0$ be constants. Set

$$\alpha(a) = \begin{cases} \alpha_N(a) \text{ if } \dfrac{|D_{i+1}|}{\sqrt{\tau_{i+1}^2 + \tau_i^2}} < c(c) \\ \alpha_C(a) \text{ if } \dfrac{|D_{i+1}|}{\sqrt{\tau_{i+1}^2 + \tau_i^2}} > c(c) \end{cases} \tag{7.25}$$

where $c(c)$ is a constant; for example $c(c) = 3$.

7.8 Examples

The simulation model of Section 7.6 is run. Observations occur at unit times. All observations have sensor standard deviation τ. The reporting cycle time is either 2 or 5 time units. The parameter indicator of change, $c(c)$ and the asset smoothing parameter if change is indicated, $\alpha_C(a)$ are varied. The TOE classification procedure (7.23) with $\tilde{\pi}_O$ modified as in (7.10) – (7.11) is used to estimate the number of units at the node. The total number of replications is 100. The summary statistics of Section 7.6 are gathered. The results appear in Table 7.2.

The results suggest that summarizing the data with an average of the observations is not as good as a weighted average with older observations having less weight. The results also suggest that using information from the previous node at the end of the previous reporting cycle is advantageous.

The physical movement and observation model of Section 7.6 is simulated again. However, there are two different sensors; one with standard deviation $\tau_L = 20$ and the other with standard deviation $\tau_H = 50$. Observations are taken at integer times and the sensor used is chosen at random. Each sensor has probability 1/2 of being chosen. If two observations are taken at a time both observations use the same sensor.

Table 7.2
Summary Statistics for Updating Procedures
With More Than One Observation Per Reporting Cycle

τ	Reporting Cycle Time	Indicator of Change $c(c)$ (7.25)	Asset Smoothing Parameter No Change $\alpha_N(a)$ (7.25)	Asset Smoothing Parameter Change $\alpha_C(a)$ (7.25)	Smooth Parameter for Number of Units $\alpha(p)$ (7.10)	Posterior Probability Used	Average of Mean Posterior Probability of Correct Number of Units (std dev. over simulation repl.)	Average of Mean Deviation of Posterior Mean Number of Assets from Correct Number (std dev. over simulation repl.)
20	2	3	1	0.1	0.1	π_O	0.87 (0.11)	32.7 (13.4)
20	2	3	1	0.1	0	π_O	0.57 (0.11)	67.6 (11.7)
20	2	3	1	0.1	0.2	π_O	0.85 (0.12)	33.5 (13.1)
20	2	3	1	0.05	0.2	π_O	0.89 (0.11)	30.50 (11.50)
20	5	3	1	0.05	0.1	π_O	0.88 (0.12)	46.7 (24.2)
20	5	3	1	0.10	0.1	π_O	0.86 (0.13)	48.6 (20.2)
20	5	2	1	0.05	0.1	π_O	0.85 (0.12)	52.1 (21.1)
20	5	3	1	1	0.1	π_O	0.70 (0.18)	71.0 (19.9)
20	5	3	0.05	0.05	0.1	π_O	0.87 (0.12)	48.2 (22.6)
50	5	3	1	0.05	0.1	π_O	0.62 (0.14)	63.2 (24.3)
50	5	3	1	0.05	0	π_O	0.46 (0.15)	80.8 (27.8)
50	5	3	1	0.05	0.2	π_O	0.63 (0.12)	64.0 (21.5)
50	5	3	1	0.1	0.1	π_O	0.62 (0.13)	66.8 (22.0)
50	5	2	1	1	0.1	π_O	0.62 (0.14)	74.5 (18.3)
50	5	3	0.05	0.05	0.1	π_O	0.60 (0.15)	65.8 (24.9)

Table 7.3 reports the simulation results. Each simulation has 100 replications. Once again, there is the indication that weighted averages of observations with older observations weighted less are better than simple averages. There is also an indication that using prior information from neighboring nodes is better than only using information from the node. The values of the various smoothing parameters do not seem to be that important.

Table 7.3
Summary Statistics for Updating Procedures
Randomly Chosen Sensors

Reporting Cycle Time	Asset Estimator Parameters			Unit Estimator Parameter	Mean Average of Posterior Probability of Correct Number of Units (std dev. over simulation repl.)	Mean Average of Deviation of Posterior Mean Number of Assets from Correct Number of Assets (std dev. over simulation repl.)
	Indicator of Change	Asset Smoothing Parameter No Change $\alpha_N(a)$	Asset Smoothing Parameter Change $\alpha_C(a)$	Smooth Parameter for Number of Units $\alpha(p)$		
5	3	1	0.1	0.1	0.79 (0.15)	58.1 (22.7)
5	3	1	0.1	0	0.55 (0.15)	88.5 (24.2)
5	3	1	1	0	0.47 (0.12)	97.1 (19.4)
5	3	1	1	0.1	0.70 (0.15)	71.9 (21.0)
5	3	0.1	0.1	0	0.57 (0.19)	77.3 (31.3)
5	3	0.1	0.1	0.1	0.76 (0.13)	55.2 (21.7)

7.9 Generalization to More Than 1 Asset Type

Suppose that a unit has J asset types. Let X_{ji} denote the i^{th} sensor observation of asset type j. Let τ_{ji} be the standard deviation of the i^{th} sensor observation of asset type j. Let t_{ji} be the time of the sensor observation and O_j be the number of sensor observations of the j^{th} asset type during the reporting cycle. A weighted average estimate of the number of assets of type j is

$$\hat{a}_j = \frac{\sum_{i=1}^{O_j} \frac{x_{ji}}{\tau_{ji}^2 \alpha(a)^{-(t-t_{ji})}}}{\sum_{i=1}^{O_j} \frac{1}{\tau_{ji}^2 \alpha(a)^{-(t-t_{ji})}}} \tag{7.26}$$

and

$$\operatorname{Var}\left[\hat{A}_j\right] = \frac{\sum_{i=1}^{O_j} \frac{\tau_{ji}^2}{\left(\tau_{ji}^2 \alpha(a)^{-(t-t_{ji})}\right)^2}}{\left(\sum_{i=1}^{O_j} \frac{1}{\tau_{ji}^2 \alpha(a)^{-(t-t_{ji})}}\right)^2}. \tag{7.27}$$

To update the probability of k units at the node at the end of the reporting cycle

$$\pi_O(k;n,t) = C\tilde{\pi}_O(k;n,t-\Delta)\prod_{j=1}^{J} \frac{1}{\sqrt{k\sigma_a(j)^2 + \operatorname{Var}\left[\hat{A}_j\right]}} \exp\left\{-\frac{1}{2}\frac{\left(\hat{a}_j - k\gamma_j\right)^2}{k\sigma_a(j)^2 + \operatorname{Var}\left[\hat{A}_j\right]}\right\} \tag{7.28}$$

where γ_j is the TOE mean number of assets of type j for 1 unit, $\sigma_a(j)$ is the TOE standard deviation of the number of assets of type j for 1 unit, C is a normalizing constant, $\tilde{\pi}_O$ is the (possibly modified) probability from the previous reporting cycle, and Δ is the length of the reporting cycle.

7.10 Updating Using Same Node Information

Consider just one asset type at a node and let $\mu(t)$ be the ground truth number of assets of that type on the node at time t. Assume observations of the asset number are taken at times $t_1 < t_2 < \ldots < t_n$ during an observation period. Let $X(t_k)$ denote the observation obtained at time t_k. Assume

$$X(t_k) = \mu(t_k) + \mathcal{E}(t_k)$$

where $\mathcal{E}(t_k)$ has a normal distribution with mean 0 and variance $\tau(t_k)^2$; the observation $X(t_k)$ could also be obtained from the binomial or multinomial sensor models of Sections 4 – 6; the appropriate variances should be used in the procedures below.

Estimates of node occupancy are obtained at the end of observation periods and use sensor observations collected during the period. Since units move, ground truth $\mu(t)$ will change as a function of t. The problem of detecting when changes occur in ground truth from observational data has received attention for many years; cf. Basseville (1988) and Lai (1995). Problems of this type are called changepoint problems. If a change in ground truth occurs during an observation period, then combining observations as proposed in Appendix A may not be the best procedure.

There will be many nodes and arcs in a typical theater-level model. Sensor observations may occur at many nodes and more than once at a node during an observation period. As a result, the statistical calculation to estimate the number of assets at a node will be done many times for each observation period. Thus, one must be careful with regard to the computational burden of the statistical calculation. Another procedure for combining observations is proposed below.

A proposed procedure for combining observations which may be from distributions having different means

We will take a generic observation period to be one time unit in length; e.g. [0,1]. Let c be a tuning parameter chosen by the analyst; good values for c are 2 and 3.

Suppose observations are taken at times $0 < t_1 < t_2 < \ldots < t_n < 1$.

1. Compute

$$D_n = \frac{\left|X(t_n) - X(t_{n-1})\right|}{\sqrt{\tau^2(t_n) + \tau^2(t_{n-1})}};$$

$$D_{n-1} = \frac{\left|X(t_{n-1}) - X(t_{n-2})\right|}{\sqrt{\tau^2(t_{n-1}) + \tau^2(t_{n-2})}}; \tag{7.29}$$

$$\vdots$$

$$D_j = \frac{\left|X(t_j) - X(t_{j-1})\right|}{\sqrt{\tau^2(t_j) + \tau^2(t_{j-1})}}.$$

2. If $\max\limits_{j=n,\ldots,2}(D_j) < c$, then compute

$$\hat{\mu} = \frac{\sum\limits_k \frac{X(t_k)}{\tau(t_k)^2}}{\sum\limits_k \frac{1}{\tau(t_k)^2}} \quad \text{and} \quad \hat{v} = \frac{1}{\sum\limits_k \frac{1}{\tau(t_k)^2}} \tag{7.30}$$

Use $\hat{\mu}$ as the estimate of the number of assets at the node and $\hat{v}$ as its variance.

3. If $\max\limits_{j=n,\ldots,2}(D_j) > c$, then find the largest j for which $D_j > c$.

$$J = \max\{j\colon D_j > c\}. \tag{7.31}$$

Compute

$$\hat{\mu} = \frac{\sum\limits_{k>J} \frac{X(t_k)}{\tau(t_k)^2}}{\sum\limits_{k>J} \frac{1}{\tau(t_k)^2}} \quad \text{and} \quad \hat{v} = \frac{1}{\sum\limits_{k>J} \frac{1}{\tau(t_k)^2}}. \tag{7.32}$$

7.11 Procedures for Updating Number and Type of Unit at a Node

For simplicity of discussion and notation we will assume there is one type of unit with only one type of asset. Let $K(n, t)$ be the number of units that occupy node n at time t and let $A(n, t)$ be the number of assets at node n at time t; n could also be an arc. Assume the prior distribution for the number of assets given the number of units with no sensor observations is

$$P\left\{A(n,0) \in da \middle| K(n,0) = k\right\} = \frac{1}{\sqrt{2\pi}\sqrt{k}\sigma_a} \exp\left\{-\frac{1}{2}\frac{(a-k\gamma)^2}{k\sigma_a^2}\right\} \tag{7.33}$$

where γ is the TOE (Table of Organization and Equipment) value for the amount of asset for one unit.

Let

$$\pi(k;n,t) = P\left\{K(n,t) = k \middle| \text{ all sensor observations during } [0,t]\right\} \tag{7.34}$$

and let

$$\begin{aligned} dF_k(a;n,t) &= P\left\{A(n,t) \in da \middle| \text{ all sensor observations during } [0,t], K(n,t) = k\right\} \\ &\equiv \frac{1}{\sqrt{2\pi}\,v_k(n,t)} \exp\left\{-\frac{1}{2}\left(a - m_k(n,t)\right)^2 \Big/ v_k^2(n,t)\right\}. \end{aligned} \tag{7.35}$$

Results from a small-scale simulation reported in Section 7.6 suggest that modifying the prior distribution used in the Bayesian updating of the distribution of the number of units at the node can improve the performance of the updating procedure. The modification uses information concerning the posterior distribution of the number and types of units at the nodes which are direct neighbors of n; the specification of direct neighbors can depend on the length of time between sensor observations.

Suppose that node n has only 2 neighbors n_- and n_+; these neighbors could be adjacent arcs. Modify the prior distribution of the number and type of units at node n at the beginning of the t^{th} observation cycle as follows.

$$\tilde{\pi}(k,n,t)=\frac{\alpha}{4}\pi(k,n_{-},t)+(1-\alpha)\pi(k,n,t)+\frac{\alpha}{4}\pi(k;n_{+},t)+\frac{\alpha}{4}\varepsilon_0(k)+\frac{\alpha}{4}\beta(k) \qquad (7.36)$$

where α is a tuning parameter; might try $\alpha=0.1$; $\varepsilon_0(k)=0$ if $k\neq 0$ and $\varepsilon_0(0)=1$; and β assigns equal probability to each possible number and type of unit combination.

Let $\hat{A}$ be the estimate of the number of assets at node at the end of the t^{th} observation cycle; let $\hat{V}$ be the variance of $\hat{A}$. The posterior distribution of the number and type of units at node n at the end of the t^{th} observation cycle is

$$\pi(k,n,t+1)=C\tilde{\pi}(k,n,t+1)\frac{1}{\sqrt{2\pi}\sqrt{\hat{V}+k\sigma(a)^2}}\exp\left\{-\frac{1}{2}\frac{\left(\hat{A}-k\gamma\right)^2}{\hat{V}+k\sigma(a)^2}\right\}; \qquad (7.37)$$

notice that the TOE values are used rather than a current estimate of the mean and variance of the number of assets at the node given the number of units.

Numerical computation considerations suggest replacing (7.37) by the following computation.

1. For each k, compute

$$f(k)=\log\tilde{\pi}(k,n,t+1)-\frac{1}{2}\log\left(\hat{V}+k\sigma(a)^2\right)-\frac{1}{2}\frac{\left(\hat{A}-k\gamma\right)^2}{\hat{V}+k\sigma(a)^2}.$$

2. Find

$$f_m=\max_k f(k).$$

3. Let

$$K=\left\{k:\left|f(k)-f_m\right|<M\right\}$$

the set of all number and types of units whose f value is within M of the maximum; the value of M can be chosen by the analyst; $M=4$ might be reasonable.

4. For $k\in K$, let

$$\pi(k,n,t+1)=\frac{\exp\{f(k)\}}{\sum_{j\in K}\exp\{f(j)\}}.$$

For $k \notin K$

$$\pi(k,n,t+1)=0.$$

7.12 Updating Using Same Node Information with Binomial-like Sensors

Consider just one asset type at a node. Let $\mu(t)$ be the ground truth number of assets of that type on the node at time t. Assume observations of the asset number are taken at times $t_1 < t_2 < \ldots < t_K$ during an observation period. Let $X(t_k)$ denote the observation obtained at time t_k. Assume the conditional distribution of $X(t_k)$ given $\mu(t_k)$ is binomial with $\mu(t_k)$ trials and probability of success $p(t_k)$. In Section 4, an estimate of $\mu(t_k)$ is proposed in (4.19)

$$\hat{\mu}(t_k)=\frac{X(t_k)}{p(t_k)}.$$

It is suggested that the distribution of $\hat{\mu}(t_k)$ be approximated by a normal distribution with variance $\tau^2(t_k) = X(t_k)(1-p(t_k))/p(t_k)$; if $X(t_k)=0$, then set $\tau(t_k)^2 = v_0^2 > 0$ where v_0^2 is chosen by the analyst. A modification to the procedure proposed in Section 7.10 is as follows. Let c be a tuning parameter chosen by the analyst.

1. Compute

$$D_k=\frac{\left|\frac{X(t_k)}{p(t_k)}-\frac{X(t_{k-1})}{p(t_{k-1})}\right|}{\sqrt{\tau^2(t_k)+\tau^2(t_{k-1})}}$$

$$D_{k-1}=\frac{\left|\frac{X(t_{k-1})}{p(t_{k-1})}-\frac{X(t_{k-2})}{p(t_{k-2})}\right|}{\sqrt{\tau^2(t_{k-1})+\tau^2(t_{k-2})}},\ldots$$

2. If $\max_{j=K,\ldots,2}(D_j) < c$, then compute

$$\hat{\mu} = \frac{\sum_k \frac{X(t_k)}{p(t_k)} \frac{1}{\tau(t_k)^2}}{\sum_k \frac{1}{\tau(t_k)^2}}$$

and

$$\hat{v} = \frac{1}{\sum_k \frac{1}{\tau(t_k)^2}}.$$

Use $\hat{\mu}$ as the estimates of the number of assets at the node and $\hat{v}$ as its variance.

3. If $\max_{j=K,\ldots,2}(D_j) > c$, then find the largest j for which $D_j > c$.

$$J = \max\{j: D_j > c\}.$$

Compute

$$\hat{\mu} = \frac{\sum_{k>J} \frac{X(t_k)}{p(t_k)} \frac{1}{\tau(t_k)^2}}{\sum_{k>J} \frac{1}{\tau(t_k)^2}}.$$

and

$$\hat{v} = \frac{1}{\sum_{k>J} \frac{1}{\tau(t_k)^2}}.$$

8. Ground Course-Of-Action Realization and Perception

8.1 The Problem

A theater-level action, in simplest form, involves the movement of units of a force, here called Blue (B), such as infantry or armor brigades or divisions, from their initial locations (network nodes) to a final destination (possibly a single network node). The ground forces may be supported by air reconnaissance and strikes. The course of action (COA) plan could be that the units arrive at the destination, D, at a particular time and in coordination, although quite possibly along different routes or paths. In reality, times of transit from place to place (node to node) may vary, e.g. for reasons of weather or natural terrain obstacles or equipment breakdown, etc., so that times of transit from node to node may vary randomly. The arrival times at, and directions of approach to, the destination may vary as well, but by design of B-force commanders. Of course the rate of advance and force allocations depend upon B perceptions concerning the opposition's actions and capabilities.

From the point of view of the opposition, here called Red (R), the *options* for a B action are known broadly. As the campaign develops the actual COA being carried out by B will become increasingly evident to R. Of course R has his own intended COA, which may develop spatially in conformity with what his C3I system tells him about B.

Given the above general background we wish to provide an analyst with the capability of obtaining and maintaining a quantitative perception of one opponent's COA (here B) by the other (here R), based on the latter's sensor information and prior probabilities of the various COAs. The methodology applies to either side.

8.2. Routes and Corridors of Advance as Part of a Course of Action

The definition of a COA involves a *mobility corridor*, which may comprise several specific *routes*, e.g. highways or trails and neighboring territory over a broad geographical area; it may even involve a sea-land amphibious landing action. Only one route may be used by all units in the particular COA, or several routes may be used by different numbers of units. It is assumed that elements of major forces such as brigades or divisions (units) move on such defined routes, i.e. along one or more *axes of advance* at a speed influenced by the type of unit and the nature of the route. It is a convenient modeling assumption that units are detected on such routes; the probability of detection at any time along a route (an arc, or node, in the current model) depends on the unit sizes and types (their asset portfolio) occupying the route at the time and the types of surveillance assets (sensors) viewing the route at the time.

Consider the problem of detecting and classifying an advancing force. It seems reasonable to think in terms of sensor updating at regular time intervals Δ apart, where $\Delta = 6$ to 12 hours when thinking about ground unit advance. Suppose no units have been detected up to time $k\Delta$ on the routes within a given corridor. The observing force must decide where to place his reconnaissance effort within the particular corridor for the following time period $(k\Delta, (k+1)\Delta]$, and he must update his probability of corridor usage depending upon reports received from his reconnaissance system within that time period. One way of proceeding is to specify that particular units will use particular routes within the corridor, advancing at specific speeds. If they start the journey at specified times then it can be forecast where they will be at specified times in the future. This implies that times of exposure on particular parts of a particular route during time interval $(k\Delta, (k+1)\Delta]$ are specified. If reconnaissance assets are assigned to those locations during those times the only reason that the units will not be detected is through failure of the assets to perform adequately, given the presence of the units sought.

Uncertainties

Inherent in the above situation are various realistic variabilities and uncertainties. First, the exposure times of particular parts of a particular route (arc segments and node residence times in the arc-node model) will vary because start times and transit times along route segments (arcs, nodes) will tend to vary, and are actually modeled as varying randomly. Note that the probability law of transit and residence time variation that governs movement is supposed known to the agent performing surveillance; this is optimistic and can be changed. It may be that the maneuvering force will deliberately increase the variability of its movement so as to throw off the enemy reconnaissance effort.

Second, uncertainty exists concerning reconnaissance or surveillance asset effectiveness: successful detection performance is not certain, even if assets are employed during a maneuvering unit's exposure time in a particular geographical region. The probability that the reconnaissance asset flies close enough to the force to detect *etc.*, is not unity; the probability that overhead surveillance assets detect may not be unity because of cloud cover or weather conditions. Realistically, too, delays will exist in processing detection reports and in corroborating them.

A third uncertainty factor exists concerning the size and asset composition (e.g. tanks, APCs, infantry) of the maneuvering units.

8.3 Analyst Tasks and View

In what follows we describe the way in which an analyst must proceed to use J-STOCHWARS; we attempt to describe this stepwise, although there probably are gaps that must be filled. In particular we focus on B's ground force movement, and on R's perception of it by reconnaissance assets that are mainly airborne or in space, until Blue's objective is reached, where ground forces are assumed to be present.

Analyst Steps: Modeling Red's Perception of Blue's COA.

(1) Establish Spatial Network: physical nodes and connecting arcs

(2) Specify current objective for (e.g.) Blue and a schedule by which forces (specify types) move from Source nodes to (current) Destination.
Also for opposition, Red.

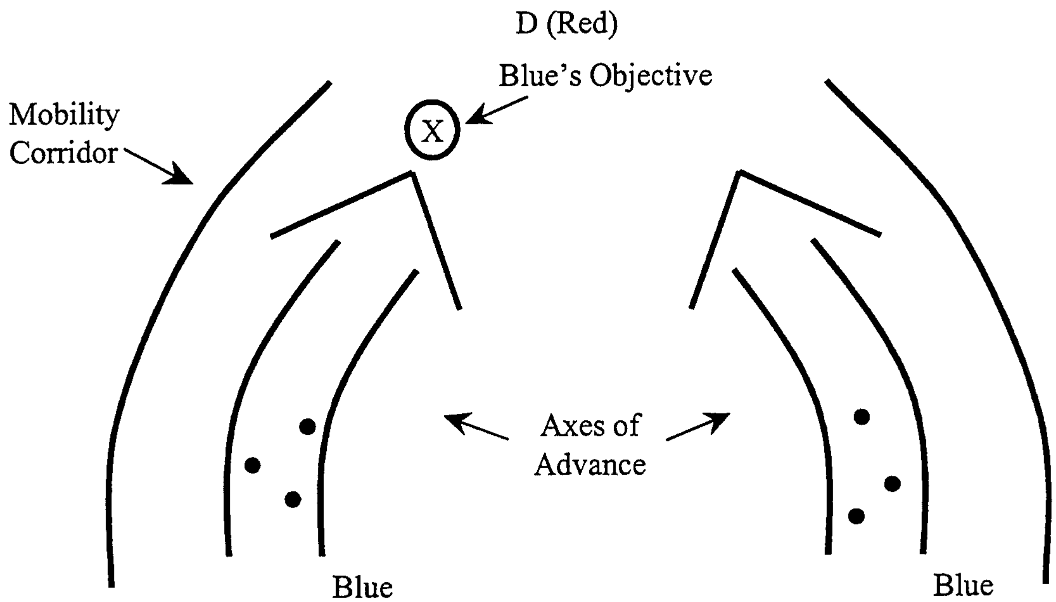


Figure 8.1

(3) Initial/Blue COA Determination: analyst specifies geography of the advance, options, and the numbers and timing of units advancing, first without opposition (certain forms of opposition can probably be optionally automatically represented, e.g. mine fields or even airstrikes). Broad *mobility corridors* will be set up for B, within which more specific *axes of advance* are specified. Red (R) must, in turn, assess B's *avenues of approach*. The model first takes these to be the same as the axes of advance. The B units will advance along arc-node paths within these.

   (a) Specify nodes & arcs in each axis of advance (see above). (Tentative; analyst intervention).

   (b) Determine spatial realizations of each a.a. (avenue of approach): define possible *routes* for forces within bands defining mobility corridors. Develop alternatives within bands.

      (b-1) Routes and force types are not independent: some routes are appropriate only for certain force types.

      (b-2) Optional routes, and forces thereon, could be specified by analyst. Consider location of opposition in picking routes and timing advance.

      (b-3) Optional routes can possibly be specified "automatically", e.g., by randomization with probabilities specified by the analyst over particular route options. The analyst, acting for Red, must first develop a portfolio of routes with possible unit types on them. These are Blue COA options.

(c) (Blue) picks *one* COA option that she will actually use; i.e., a bundle of paths/routes from each movement band consistent with a COA. These can be called *sub-COAs*. In general there can be several sub-COAs, i.e. alternative routes and forces thereon within a given COA.

(c-1) Respects route-asset compatibility constraints.

(c-2) Done by analyst intervention or by randomization.

(c-3) Blue remains on route until he encounters, or anticipates, opposition. At some point may change COA. The present discussion does not yet include encounters with opposition.

(d) All of above is done on the basis of initial perception of opposition's (Red's) locations and force composition and strength. It is subject to change. Success depends on the quality/accuracy of that perception.

(4) The Blue COA option is now simulated: transit times on arcs are simulated from arc-time distributions.

(a) Detection times by R sensor assets on arcs are simulated, potentially providing information on location and type of B units and assets thereof to R; detections occur at an exponential rate during units' exposure time to the sensors; the exponential rate will be a function of sensor allocation, detection range, etc.; time-dependence is allowed by letting detection rate on an arc depend on the time.

(5) Next develop quantitative (Red) perception of the opponent's (B's) COA. This develops from sensor data and prior information.

(a) (Red) allocates sensor effort to arcs/nodes in (Blue) COA's. *Options*:

(a-1) Analyst specifies (using experience), with the aid of the Air Model. Translate into detection rates of various units/assets on various arcs.

(a-2) Automatically determined using an optimization model (an option for the future).

8.4 A Proposed COA Perception Update Procedure Based on Unit Detection.

In this section we detail an approach to COA quantitative perception updating. The approach involves simulation in order to avoid practically infeasible tailor-made analytical calculations. We first describe procedures that update perception based on whether or not units are detected. Analytical calculations are computationally difficult because COAs are not specified by single paths, but rather by mobility corridors which can contain many possible paths. For a realistic network, the number of different ways units may traverse a network for one course of action will be large. Furthermore, paths belonging to different COAs may overlap during parts of the scenarios. Hence an analytic approach to updating perception based on enumeration of all such paths does not appear practical.

(1) Initial Steps.

(a) The analyst determines a COA perception update cycle time, Δ.

(b) At times $k\Delta$ ($k = 1, 2, \ldots$) measured from an initial time (e.g. action beginning) the probabilities of the opponent's (here B's) COAs are calculated; those at time $(k + 1)\Delta$ are obtained from those at time $k\Delta$, augmented by information obtained in the time interval $(k\Delta, (k + 1)\Delta]$. In practice Δ might be 6–12 hours in duration.

(c) Let C_i be the event that the i^{th} ($i = 1, 2, \ldots, I$) COA is being pursued (by B); let $\widetilde{C}_i(k)$ be the event that C_i is perceived (here by R) to be in progress at time k. Then let $\prod_i(k) = P(\widetilde{C}_i(k))$, the posterior distribution for various COAs at time $k\Delta$. This quantifies R's current (time k) perception of B's forces and COA.

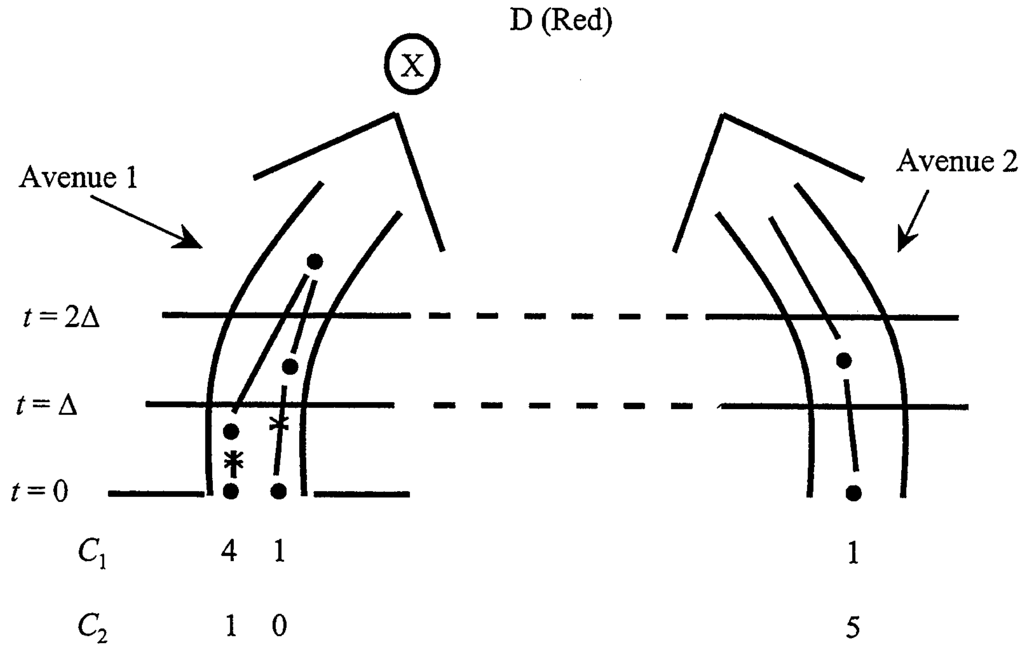


Figure 8.2

(d) Consider a given COA, so let C_i begin at time 0. Suppose for illustration we have the two alternatives shown in Figure 8.2 and that C_1 is actually being pursued. C_1 has 4 units traveling along the leftmost route in Avenue 1, 1 unit traveling along the rightmost route in Avenue 1 and 1 unit traveling along Avenue 2. At time $d_{11} \in [0, \Delta)$ a detection occurs on the left-most route in Avenue 1 (denoted A_{11}); at time $d_{12} \in [0, \Delta)$ a detection is made on A_{12}; no detections are made on A_{21}, the only path in A_2.

The above corresponds to actual data obtained in a real operation. The objective is to obtain from it a *probability* that C_1 is actually under way, plus probabilities of alternatives (here only C_2).

(2) Evaluating the Likelihood of a COA.

We want to evaluate the *probability* of observing what has been observed in terms of the possible COAs, i.e. the likelihood that one or the other COA is in force. This can then be combined with $\Pi_i(k)$ to update to $\Pi_i(k+1)$. We want to do so economically, so we choose to approximate and simplify by removing the specificity of the observations mentioned in (1),(d): we report seeing 2 detections in Avenue 1 and 0 in Avenue 2.

We now calculate the likelihood of C_1 and C_2 on the basis of the above data presentation. Here are two procedures. An exposure time during $(k\Delta, (k+1)\Delta]$ is the amount of time in $(k\Delta, (k+1)\Delta]$ that units on an arc are subject to sensor observation; it is a function of transit times, the type of sensor, and environmental variables. Let $\lambda_{11}(k)$ be the exponential rate of detection for one unit during its exposure time on A_{11} in $(k\Delta, (k+1)\Delta]$. The following is a simulation approach to evaluating the likelihood of a COA.

(a) **Simulation**

(a-1) Suppose C_1 is in effect. Simulate the exposure times, on each arc, of 4 units on A_{11} and 1 unit on A_{12} during $(0, \Delta]$, and also simulate for each exposure time the number of times detections occur. Note that the detection rate on A_{11} is $\lambda_{11}(k)4$ and on A_{12} is $\lambda_{12}(k)1$ for C_1; it is $\lambda_{11}(k)1$ for C_2.

Use the constraint that if a detection occurs on an arc in an avenue of approach no more can occur thereafter on that arc or on routes/paths that continue it (unless the arc ends at a node out of which several arcs occur; in that case the units originally detected could be lost, and might need to be re-detected). There can thus be

n_{00} replications with no detections on either A_{11} or A_{12};
n_{10} replications with 1 on A_{11}, 0 on A_{12};
n_{01} replications with 0 on A_{11}, 1 on A_{12};
n_{11} replications with 1 on A_{11}, 1 on A_{12};

(a-2) To estimate the likelihood of C_1 given 2 detections on A_1 (one on A_{11} and one on A_{12}) quote $\left(\frac{n_{11}(1)}{n_t(1)}\right)^1$; here $n_{11}(1)$ refers to one detection on each path for A_1 and $n_t(1)$ is the number of simulation replications.

Presumably one also simulates events on A_2, using $\lambda_2 \cdot 1$ for the detection rate for C_1; the probability of the observed data is $\frac{n_0(2)}{n_t(2)}$ where $n_0(2)$ is the number of $n_t(2)$ replications that result in 0 detections on A_2.

Hence the (combined, estimated) likelihood for C_1 is $\alpha(C_1; \text{data}) = \left(\frac{n_{11}(1)}{n_t(1)}\right)\left(\frac{n_0(2)}{n_t(2)}\right)$.

(a-3) Next re-simulate the detections but with the detection rates appropriate to COA C_2

C_2: $\lambda_{11} \cdot 1$ on A_{11} (and A_{12});

$\lambda_2 \cdot 5$ on A_2

and tabulate the same simulated ratios. Apply Bayes to update.

(a-4) The above procedure will work for arbitrarily complicated setups, but may be computer intensive. An alternative is

(b) **Hybrid Simulation and Calculation of Likelihood.**

It is advantageous to avoid the simulation of detections in the process of updating perceptions of alternative COA probabilities. A possible way of doing so is as follows.

(b-1) Consider the r^{th} exposure time realization for arcs and nodes occupied during $(k\Delta, (k+1)\Delta]$. Let

$$\bar{\lambda}_i(k;r) = \sum_{j \in J(k)} \lambda_j(k) e(k;r)_{ji}\, U_{ji}(k)/\Delta \tag{8.1}$$

be the mean detection rate over the entire mobility corridor for period k: times $(k\Delta, (k+1)\Delta]$. In (8.1) we sum up the basic detection rates on all arcs and nodes that are occupied (have non-zero exposure times $e(k; r)_{ji}$) during period k); these are weighted by the appropriate exposure times for COA C_i, and by the arc/node loads (unit sizes), $U_{ji}(k)$ that appear on arc/node j for COA C_i.

We now compute that the probability of 0 detections as

$$P\{\boldsymbol{D}(k) = 0 | C_i, r\} = e^{-\bar{\lambda}_i(k;r)\Delta}$$

This is the likelihood element for C_i given the exposure times of realization r; if exposure times are in effect sampled independently we use

$$\bar{p}_0(i,k) = \frac{1}{R}\sum_{r=1}^{R} e^{-\bar{\lambda}_i(k;r)\Delta}$$

where R is the number of replications.

(b-2) Suppose at least one detection takes place, an event of probability $1-e^{-\bar{\lambda}_i(k;r)\Delta} \cong \bar{\lambda}_i(k;r)\Delta$ if $\bar{\lambda}_i\Delta$ is small, as may be true in practice.

Simple Approximate Approach When Detections Occur in $(k\Delta, (k+1)\Delta]$.

Assume that if 1 or more detections occur in the k^{th} interval

$$P\{\mathbf{D}(k) = d(k) | C_i, r\} = e^{-\bar{\lambda}_i(k;r)\Delta} \frac{(\bar{\lambda}_i(k;r)\Delta)^{d(k)}}{d(k)!}.$$

If the above is averaged over the R replications of exposure time we arrive at an expression for

$$P\{\mathbf{D}(k) = d(k) | C_i\} \cong \bar{p}_{d(k)}(i,k)$$

an average of Poisson probabilities. This is the semi-parametric version of the Simulation approach above, in (2), (a).

A More Correct (but Difficult) Approach

The above procedure is not literally correct (although it may be an adequate approximation) since once a unit detection occurs on a specific route there are assumed to be no more detections on that route unless the unit reaches a physical node where she can be hidden or emerge "branched" or "split" into two or more unit segments (e.g. a division may divide into several brigades) which take different routes. It becomes necessary for the analyst to decide *how* the split occurs; since this can be done in several ways a multiplicity problem threatens; one way to handle this is by sampling. A path on which a detection has occurred during period k is essentially pruned *for the purpose of detection* after the detection on it in period k; this holds into period $k+1$, or until a node is reached with a split. Note: another way of pruning is to ignore all detection events (counts) greater than 1.

After detection occurs classification assets are brought into play. The updating that follows will be described subsequently.

8.5 Updating the Perception of COA Using Asset-Counting Sensors.

Once units are detected, asset counting sensors can be employed to obtain more information. Suppose units have been detected at node N_n. Let $S_1(n,j;k)$, …, $S_{b_n}(n,j;k)$

be the sensor observations of assets of type j at node N_n during the observation period $(k\Delta, (k+1)\Delta]$. We will assume Δ is small enough so that the number of assets at the node is constant over the period. The least-squares estimate of the number of assets of type j at node N_n based on the latest sensor information only (gathered during period $k\Delta, (k+1)\Delta$ is

$$\hat{A}(n,j;k)=\sum_{\ell=1}^{b_n}\frac{s_\ell(n,j;k)}{\tau_{nj}^2(\ell)}\bigg/\sum_{\ell=1}^{b_n}\frac{1}{\tau_{nj}^2(\ell)} \tag{8.2}$$

where τ_{nj}^2 is the variance of the error of the sensor observation $S_\ell(n,j;k)$.

We can interpret the results of the sensor observations as resulting in a normal distribution for the number of assets of type j at the node during the time period; the distribution has mean

$$m_{nj}(k)=\left[\sum_{\ell=1}^{b_n}s_\ell(n,j;k)/\tau_{nj}^2(\ell)\right]\bigg/\sum_{\ell=1}^{b_n}\frac{1}{\tau_{nj}^2(\ell)} \tag{8.3}$$

and variance

$$\upsilon_{nj}^2(k)=\frac{1}{\displaystyle\sum_{\ell=1}^{b_n}\frac{1}{\tau_{nj}^2(\ell)}} \tag{8.4}$$

Let $J(\alpha;k)$ be the collection of all nodes that might be occupied during the period $(k\Delta, (k+1)\Delta]$ for a particular avenue of approach α. The result of all sensor observations during $(k\Delta, (k+1)\Delta]$ is a distribution of the total number of assets of type j at all the nodes in $J(\alpha;k)$ which is normal with mean

$$\overline{m}_{\cdot j}(\alpha;k)=\sum_{N_n\in J(\alpha;k)}m_{nj}(k) \tag{8.5}$$

and variance

$$\overline{\upsilon}_{\cdot j}^2(\alpha;k)=\sum_{N_n\in J(\alpha;k)}\upsilon_{nj}^2(k); \tag{8.6}$$

how well this distribution reflects the true number of assets that are on the avenue of approach α will be affected by the number of sensor observations at each relevant node.

A particular COA c has a distribution of the total number of assets of type j at nodes $N_n \in J(\alpha; k)$ in the avenue of approach α during the time period $(k\Delta, (k+1)\Delta]$, $\overline{A}_j(\alpha; k)$,

$$P\{\overline{A}_j(\alpha;k) \in da | COA = c\} = \xi(a; \mu_j(\alpha,k;c), \sigma_j^2(\alpha,k;c))da \tag{8.7}$$

where $\xi(a; \mu, \sigma^2)$ denotes the normal density function with mean μ and variance σ^2; the mean $\mu_j(\alpha, k; c)$ is obtained from the TOEs (Table of Organization and Equipment) for the types and numbers of units using the avenue of approach α for the COA c.

The procedure to update Red's perception of Blue's COA is as follows. Let $\Pi(c; k)$ be Red's posterior probability at time $k\Delta$ that Blue is following COA c. To obtain a posterior probability which incorporates the sensor information obtained in $(k\Delta, (k+1)\Delta]$

$$\begin{aligned}\Pi(c;k+1) &= D\Pi(c;k)\prod_{\alpha}\prod_{j}\int \xi(a,\mu_j(\alpha,k;c),\sigma_j^2(\alpha;k;c))\xi(a,\overline{m}_j(\alpha;k),\overline{\upsilon}_j^2(\alpha;k))da \\ &= D\Pi(c,k)\prod_{\alpha}\prod_{j}\xi(\overline{m}_j(\alpha;k);\mu_j(\alpha,k;c),\overline{\upsilon}_j^2(\alpha,k)+\sigma_j^2(\alpha;k,c))\end{aligned} \tag{8.8}$$

where the product α is over the avenues of approach, the product j is over the asset types and D is a normalization constant.

Note: the above can be adjusted for the time that the sensors have to "measure" the unit assets during $(k\Delta, (k+1)\Delta]$; it can also be made to incorporate a likelihood component informative of assets from the *detection*. Note also that since the procedure is based on the total numbers of assets each COA has on different avenues of approach, the procedure will produce more discrimination between COAs the more differentiated the COAs are. The COAs will be better differentiated the greater the difference of the numbers of assets on different avenues of approach for different COAs. The COAs will also be better differentiated in situations in which the avenues of approach do not have nodes and arcs in common.

8.6 Numerical Considerations

The proposed updating of the course of action of (8.8) is susceptible to numerical problems. We propose to first compute the logarithm of the normal density functions. First compute

$$\ell_1(c,k+1)=\sum_{\alpha}\sum_{j}-\frac{1}{2}\left(\overline{m}_j(\alpha;k)-\mu_j(\alpha;k,c)\right)^2\Big/\left[\overline{\upsilon}_j^2(\alpha;k)+\sigma_j^2(\alpha;k,c)\right] \qquad (8.9)$$

$$\ell_2(c,k+1)=\sum_{\alpha}\sum_{j}-\frac{1}{2}\ln\left(\overline{\upsilon}_j^2(\alpha;k)+\sigma_j^2(\alpha;k,c)\right) \qquad (8.10)$$

$$\ell(c,k+1)=\ell_1(c,k+1)+\ell_2(c,k+1)+\ln\Pi(c,k). \qquad (8.11)$$

Order $\ell(c, k+1)$ from largest to smallest; e.g.

$$\ell_{(1)}(k+1)=\max_c \ell(c,k+1); \qquad (8.12)$$

call these $\ell_{(1)}(k+1)$, $\ell_{(2)}(k+1)$, ... Let $c_{(1)}$, $c_{(2)}$ be the corresponding Course of Action (COA), e.g.

$$c_{(1)}=\left\{c\colon \ell(c,k+1)=\max_c \ell(c,k+1)\right\}; \qquad (8.13)$$

if there is more than one COA in the set, order them in some fashion.

Consider those COA's such that

$$\left|\ell_{(1)}(k+1)-\ell_{(j)}(k+1)\right|<4; \qquad (8.14)$$

call these $c_{(1)}$, $c_{(2)}$, ..., $c_{(m)}$. Let R be the number of COA's not among these m.

Let $0<\alpha\le 1$ be a parameter to be determined by the analyst; e.g. $\alpha = 0.9$.

For $c\in\left\{c_{(1)},c_{(2)},\ldots,c_{(m)}\right\}$, put

$$\Pi(c;k+1)=\alpha\frac{\exp\left\{\ell(c,k+1)-\ell_{(1)}(k+1)\right\}}{\sum_{j\in\left\{c_{(1)},\ldots,c_{(m)}\right\}}\exp\left\{\ell(j,k+1)-\ell_{(1)}(k+1)\right\}}. \qquad (8.15)$$

For $c\notin\left\{c_{(1)},c_{(2)},\ldots,c_{(m)}\right\}$, put

$$\Pi(c;k+1) = (1-\alpha)\frac{1}{R}. \tag{8.16}$$

Another possible procedure to update the probabilities of the different COA's is the following. Modify the prior

$$\tilde{\Pi}(c,k) = (1-\alpha)\ \Pi(c,k) + \alpha\frac{1}{|C|}. \tag{8.17}$$

where $|C|$ is the number of different COA's and α is a parameter chosen by the analyst, e.g., $\alpha = 0.1$. Compute

$$\ell(c,k+1) = \ell_1(c,k+1) + \ell_2(c,k+1) + \ln\tilde{\Pi}(c,k). \tag{8.18}$$

Order $\ell(c, k+1)$ from largest to smallest and define $\ell_{(j)}(k+1)$, $c_{(j)}$, as before. Let

$$\Pi(c,k+1) = \frac{\exp\{\ell(c,k+1) - \ell_{(1)}(k+1)\}}{\sum\limits_{j\in\{c_{(1)},\ldots,c_{(m)}\}} \exp\{\ell(j,k+1) - \ell_{(1)}(k+1)\}} \tag{8.19}$$

for $c \in \{c_{(1)}, c_{(2)}, \ldots, c_{(m)}\}$; for $c \notin \{c_{(1)}, \ldots, c_{(m)}\}$ let $\Pi(c, k+1) = 0$.

8.7 Degrading the Estimate of the Number of Assets on a Node for Which There Are No Recent Sensor Observations

Suppose the last sensor observation of node n occurs during observation period k. Let $\hat{A}_j(k;n)$ be the estimate of the number of assets of type j on node n at the end of observation period k; let $\upsilon_j^2(k;n)$ be the variance of the estimate. Suppose there are no sensor observations of the assets at node n for the next m observation periods. Since the node has not been observed for m observation periods, there may be more uncertainty concerning the estimate of the number of assets of type j at node n. Thus we will inflate the variance, by setting

$$\upsilon_j^2(k+m;n) = \frac{\upsilon_j^2(k,n)}{\alpha^m} \tag{8.20}$$

where $0 < \alpha \leq 1$ is a parameter determined by the analyst. If $\alpha = 1$, there is no inflation. The standard deviation is inflated by $(1/\sqrt{a}\,)^m$.

Inflation Factor for Standard Deviation of the Estimate: $1/\sqrt{a}$

$\sqrt{a}\backslash m$:	1	2	3	...	8	...	16
0.8	1.25	1.6	1.95		6.0		35.5
0.9	1.1	1.2	1.4		2.3		5.4
0.95	1.05	1.1	1.2		1.5		2.3

Suppose the observation intervals are 3 hours long. If $\sqrt{a} = 0.9$ and there have been no observations for 24 hours (8 observation periods), the standard deviation of the estimate is multiplied by 2.3. It may be appropriate for nodes and arcs to have individual α's; one α for a node/arc which units travel through; another α for a node on which assets are known to stay for a period of time.

To improve the numerical stability of the COA updating calculation, it may be advantageous to first do a test of hypothesis that the estimated number of assets of type j at node n is different from zero.

Let $J(\alpha; k)$ be the collection of all nodes that might be occupied during the observation period $[k\Delta, (k+1)\Delta]$ for a particular avenue of approach. For each node $n \in J(\alpha; k)$ and type of asset j, compute

$$B_j(k;n) = \frac{\hat{A}_j(k+1;n)}{\upsilon_j(k;n)}$$

where $\upsilon_j(k; n)$ has been multiplied by $(1/\sqrt{a}\,)^m$ if node n has not been observed by a sensor during the last m observation periods. If

$$B_j(k;n) < \beta$$

then do not include node n and asset j in the calculation for COA update where β is chosen by the analyst; $\beta = 2$ may be a good choice. If all asset estimates are 0 for $J(\alpha; k)$, let $\overline{\upsilon}_j^2$ be the smallest variance.

8.8 A Procedure to Update COA, Using Asset Counting Sensors, Which Recognizes the Number and Type of Unit on a Node/Arc Changes Over the Updating Period

The estimates of the number and type of units on each node and the number of assets of each type on a node is updated at the end of each sensor update cycle; nominally 2 hours, cf. Sections 7.6 and 7.11.

The COA update cycle length is often longer; nominally 6-12 hours.

It is proposed that the total number of assets of type j in the avenue of approach α, $\overline{A}_j(\alpha; k)$ be estimated as the sum of all the estimates of the number of assets of type j on all the nodes and arcs in avenue α obtained in the latest sensor update cycle. The estimation procedure for the number of assets of type j on a node/arc would incorporate procedures in Section 7.10 that address a changing ground truth during the sensor update cycle and that of Section 8.7 which degrades the estimate on a node/arc for which there are no recent sensor observations.

Let m_{nj} be the estimate of the number of assets of type j, on node/arc n at the end of the latest sensor update cycle and v_{nj}^2 be its variance. The estimate of the total number of assets of type j in a particular avenue of approach α is obtained from (8.5). Finally the probability that each COA is being executed is obtained from (8.9) – (8.12) and (8.17) – (8.19).

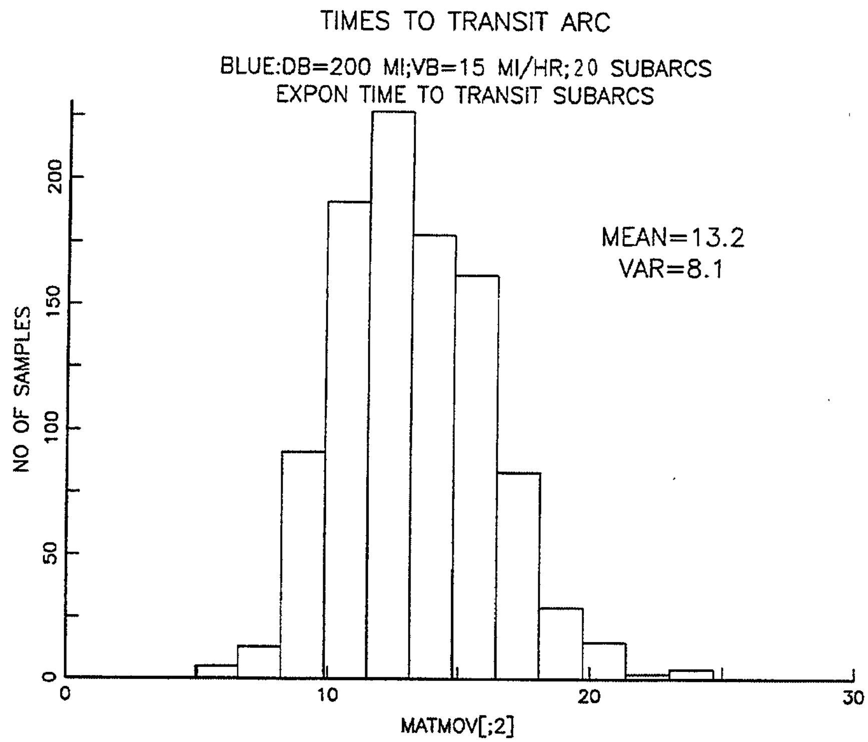


Figure 1

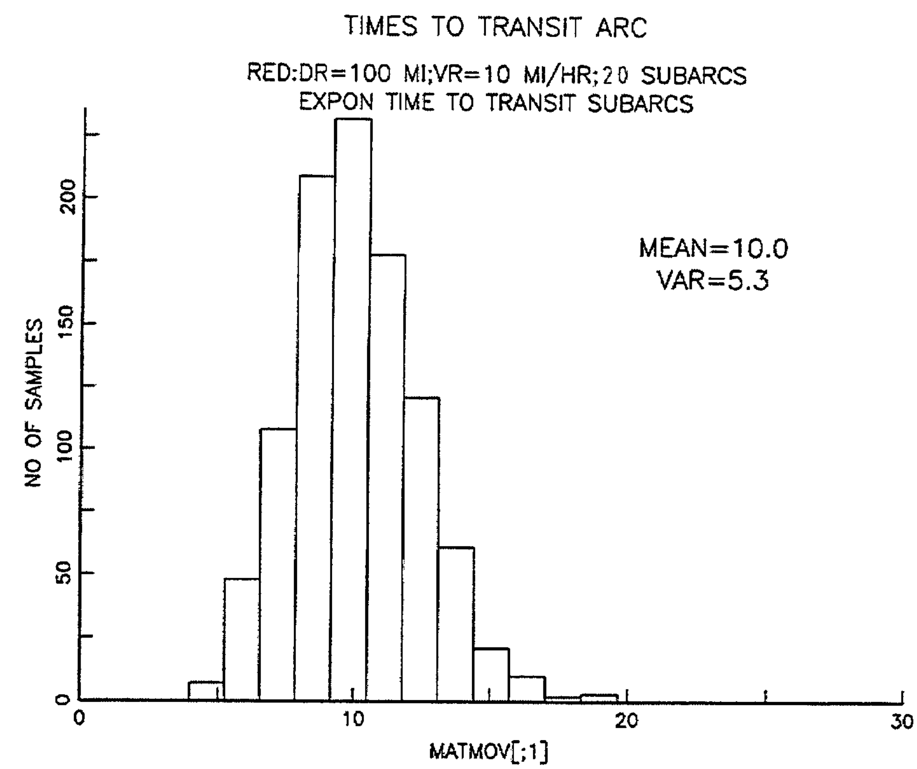


Figure 2

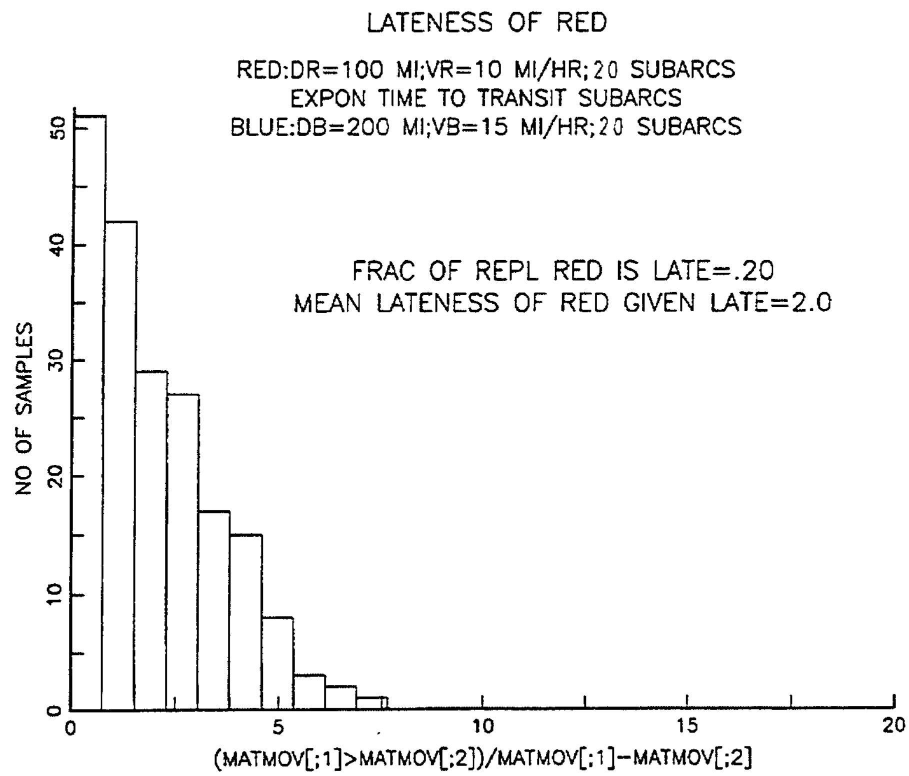


Figure 3

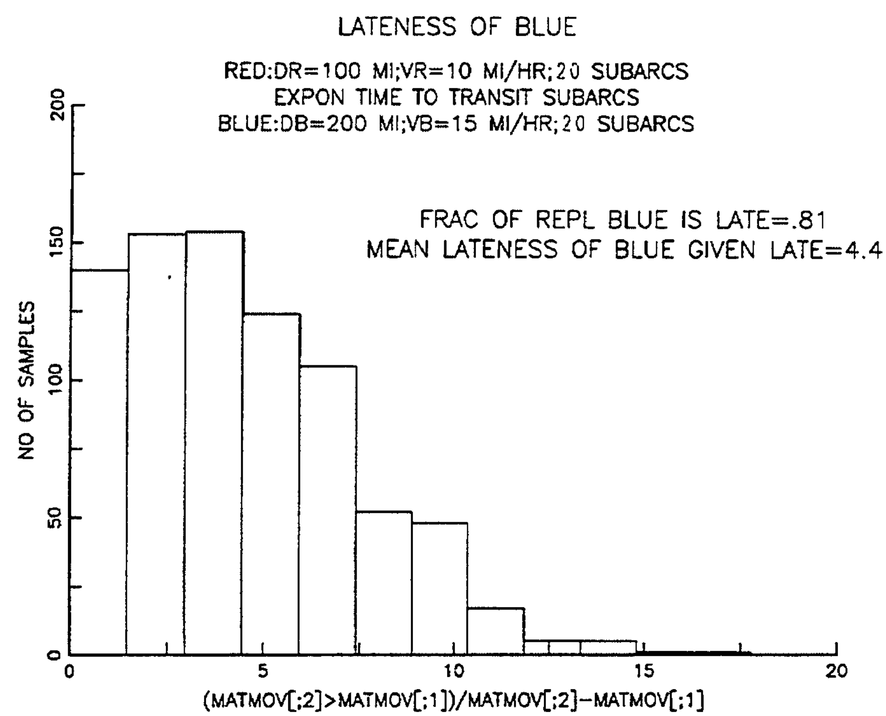


Figure 4

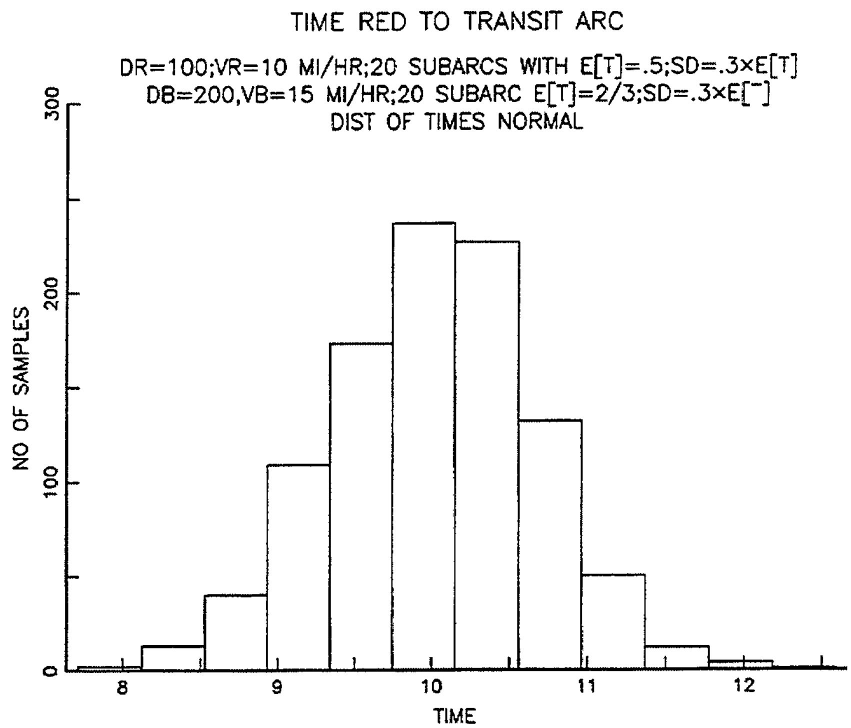


Figure 5

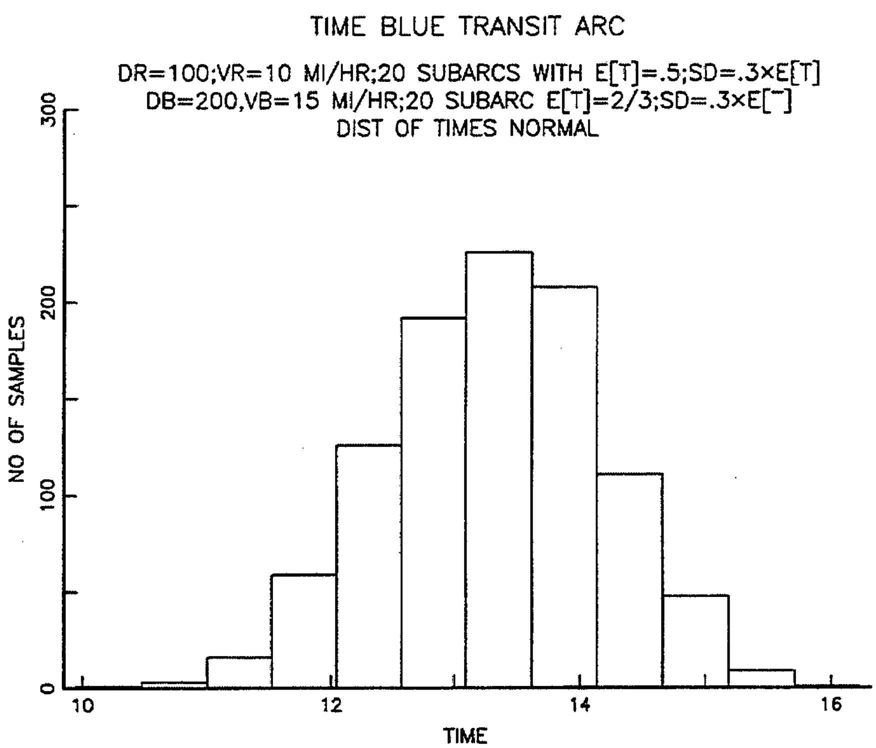


Figure 6

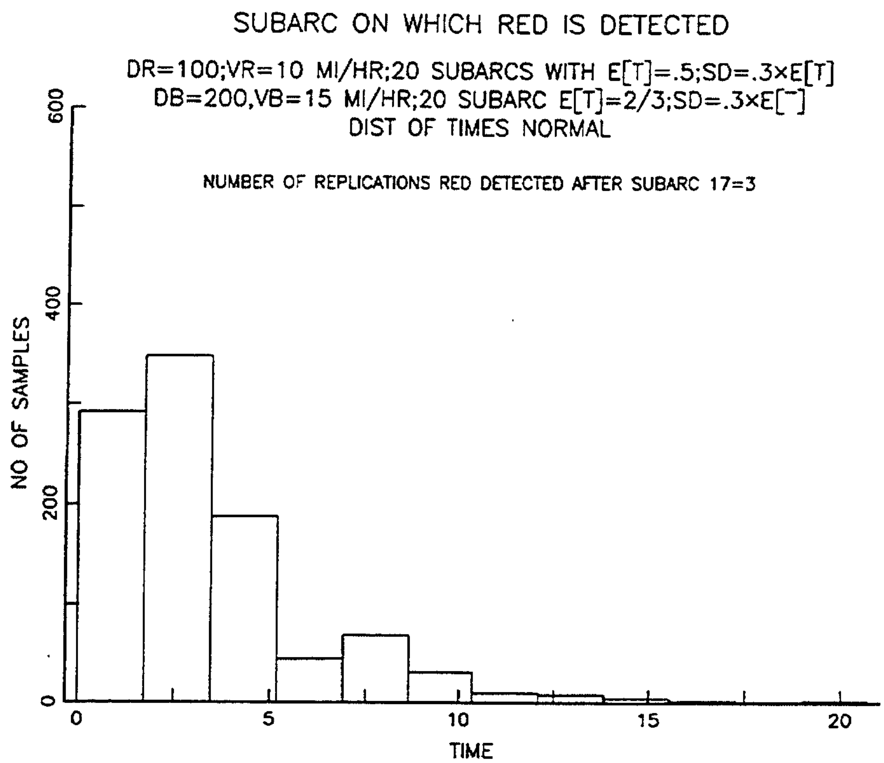


Figure 7

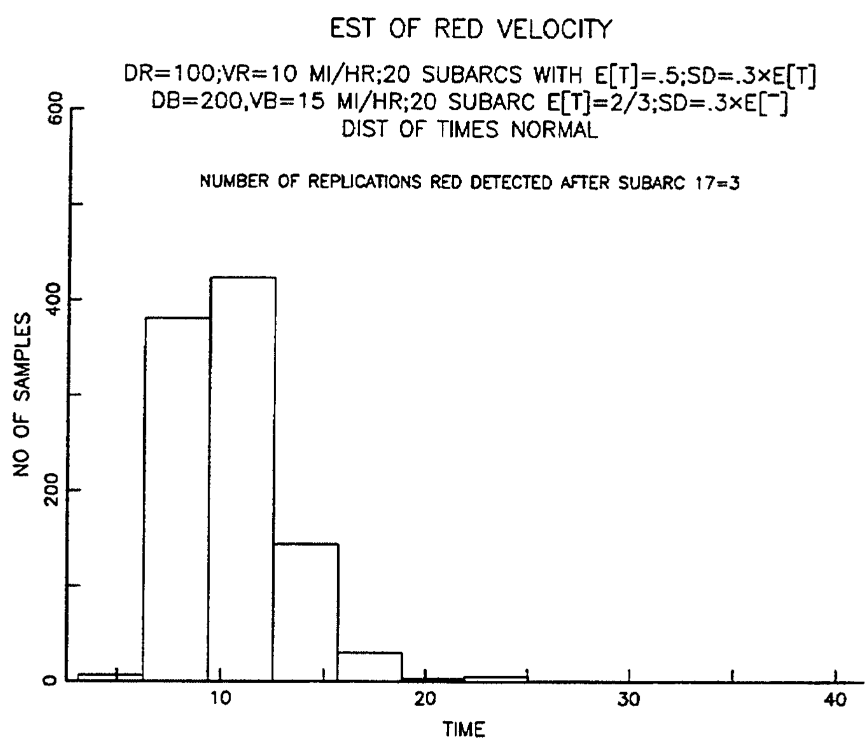


Figure 8

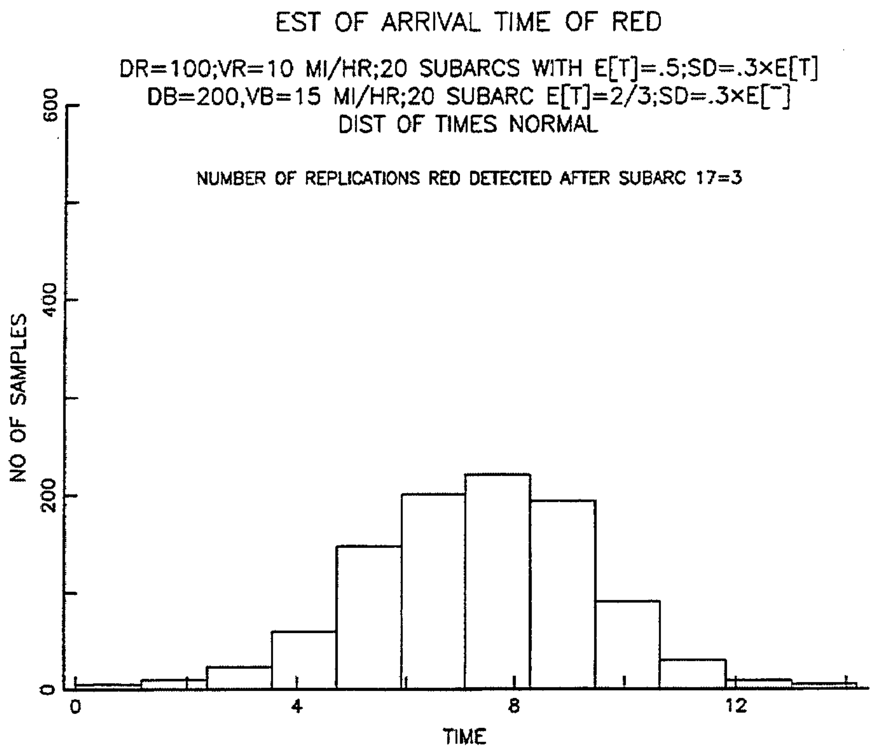


Figure 9

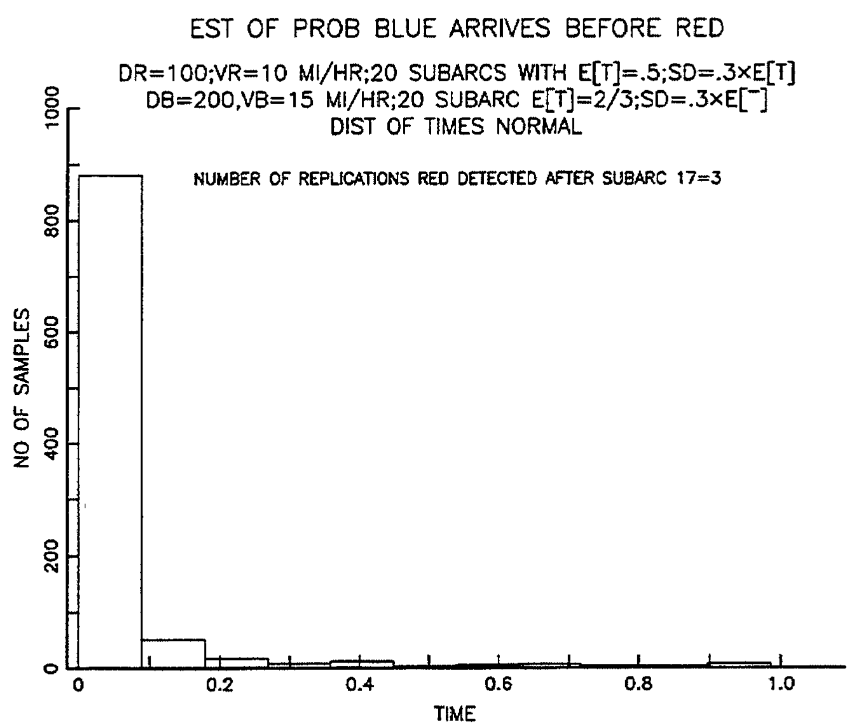


Figure 10

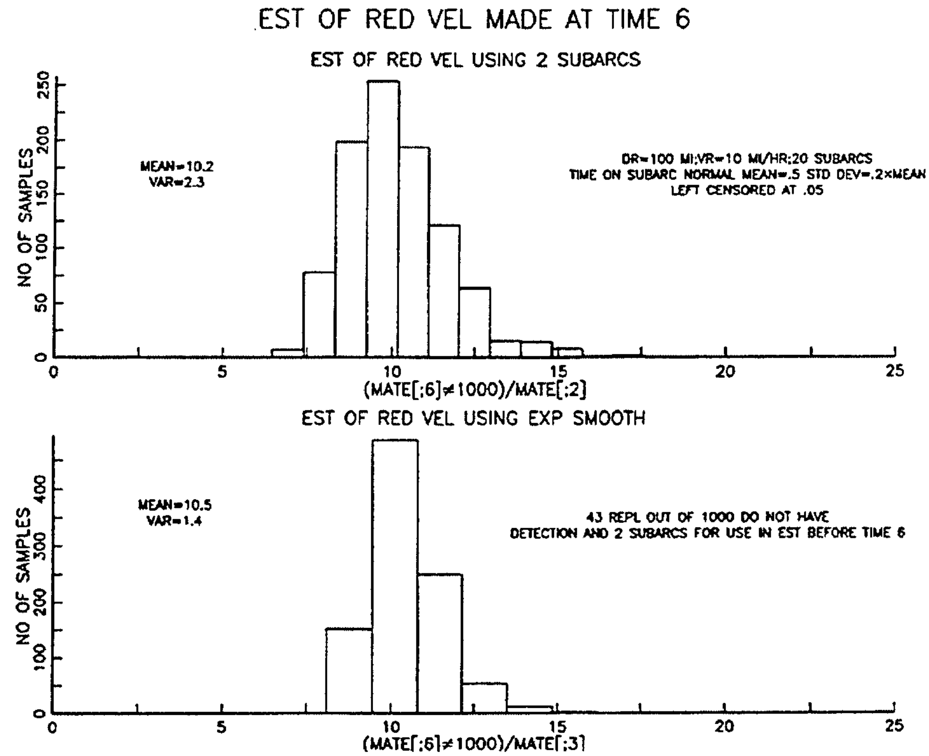


Figure 11

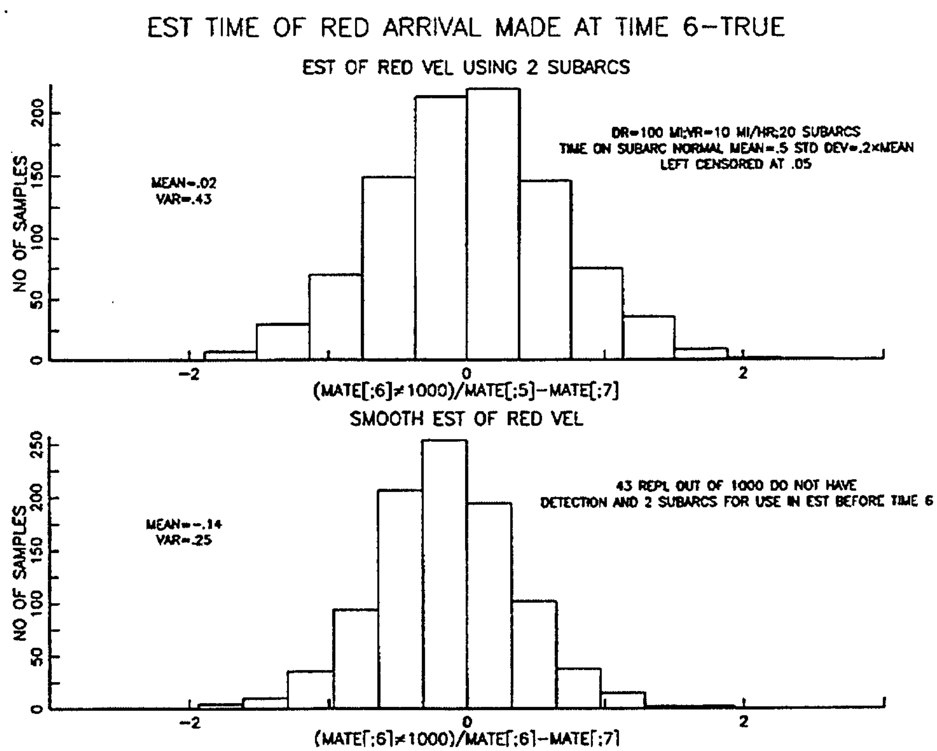


Figure 12

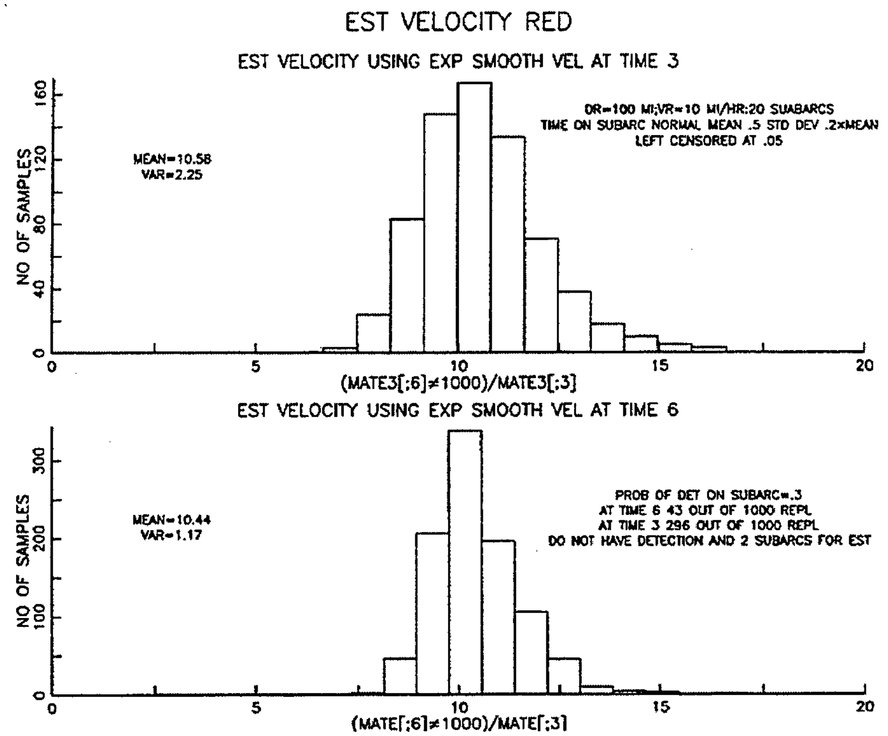


Figure 13

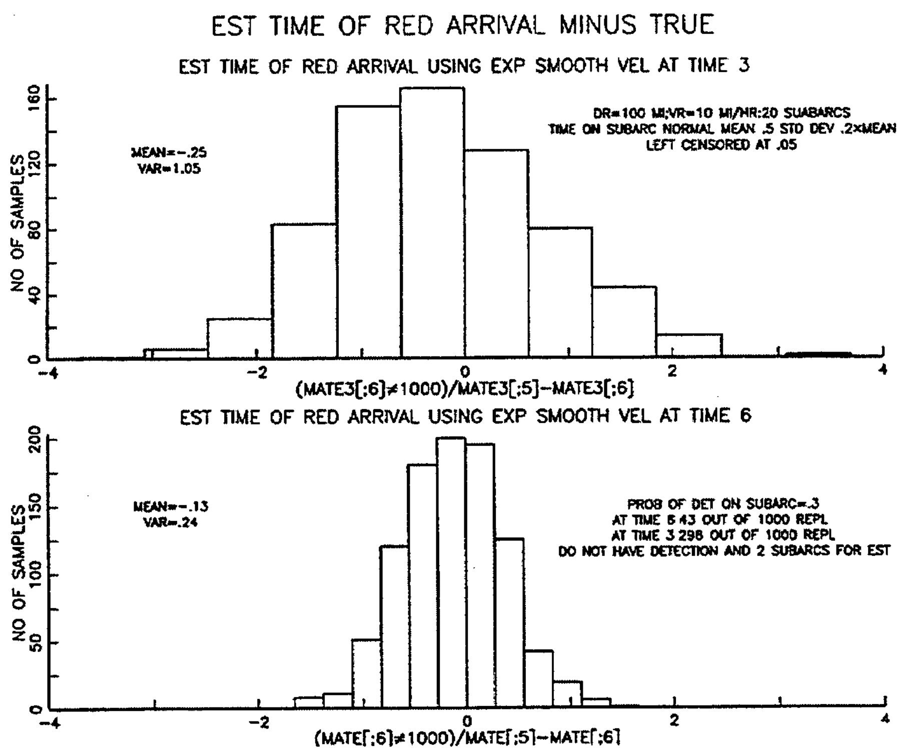


Figure 14

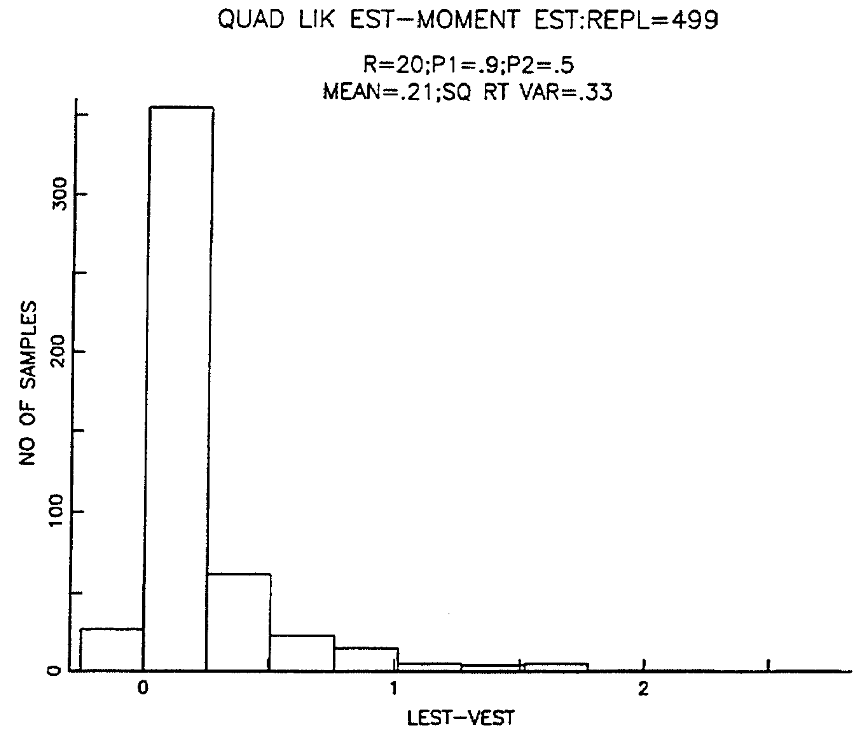


Figure 15

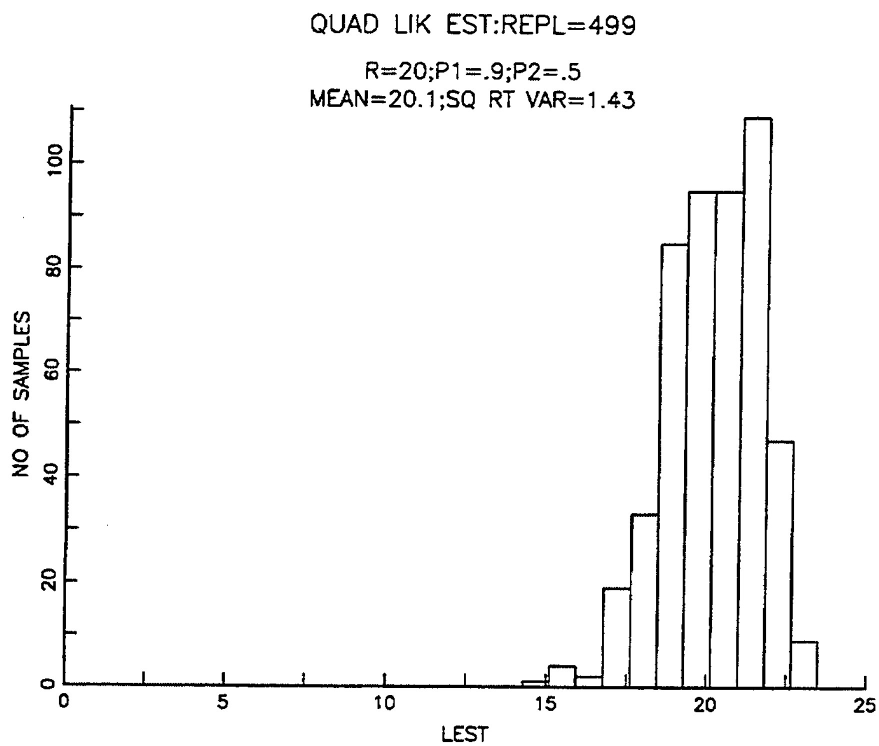


Figure 16

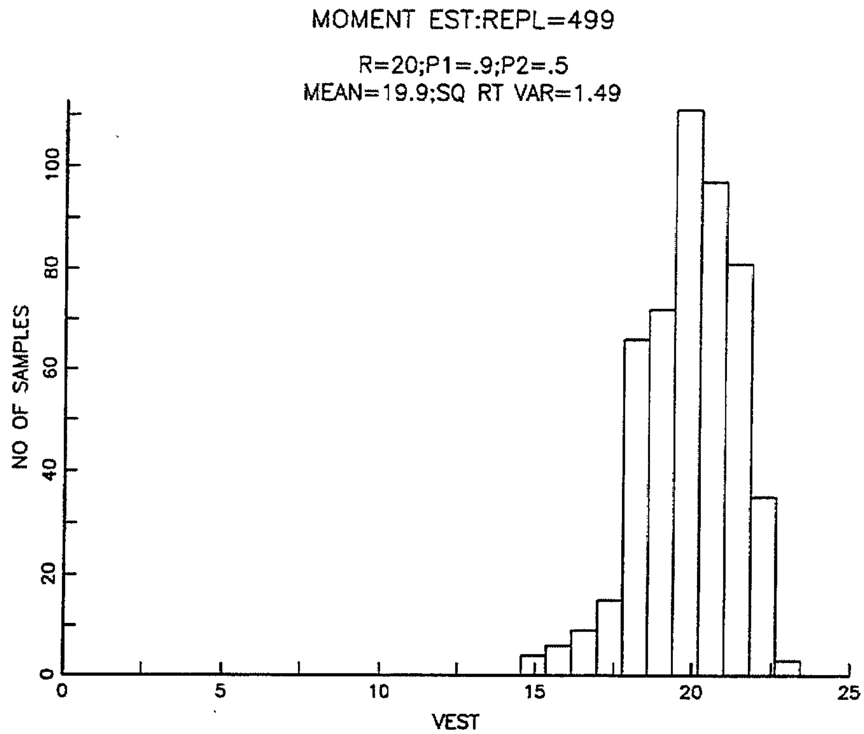


Figure 17

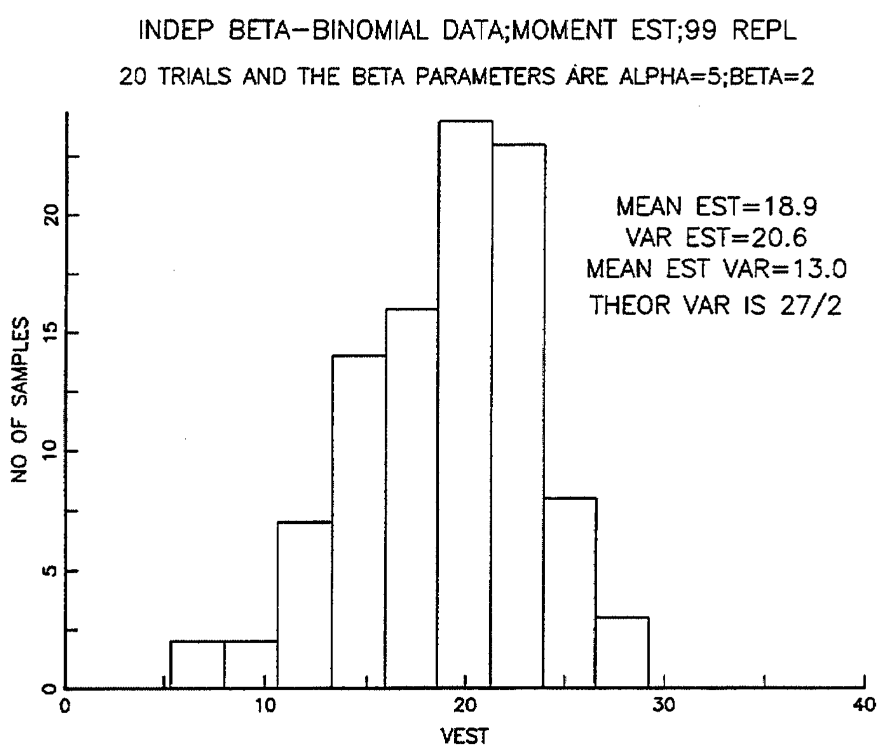


Figure 18

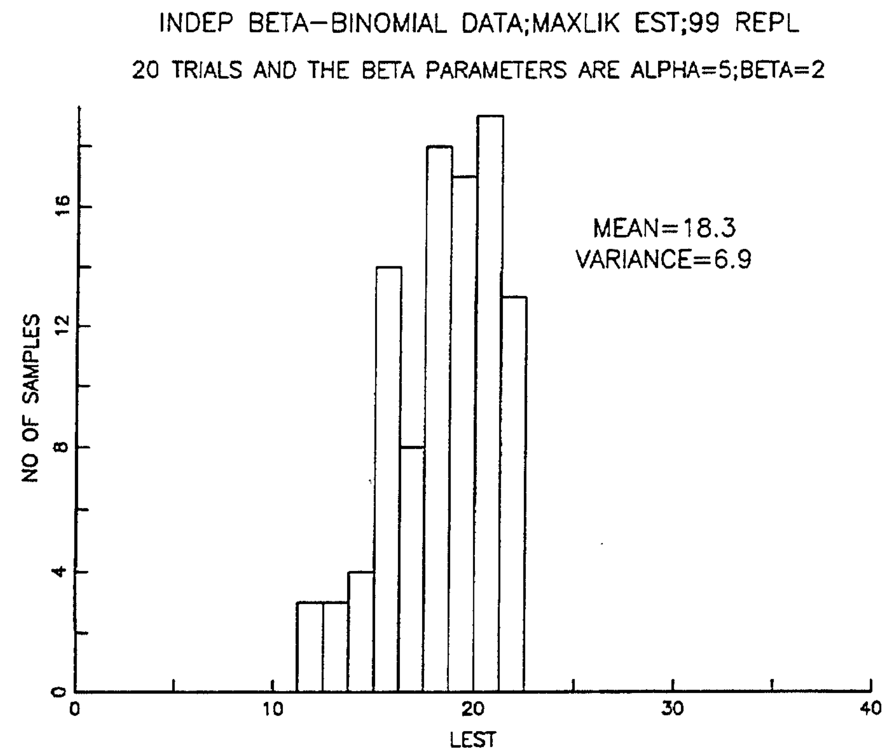


Figure 19

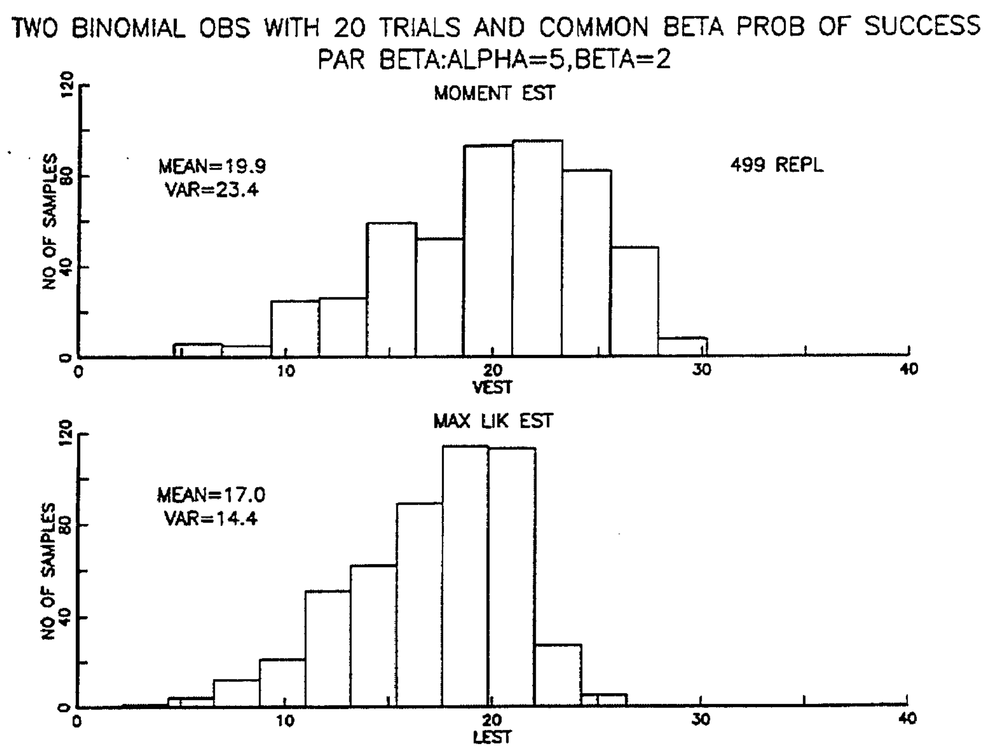


Figure 20

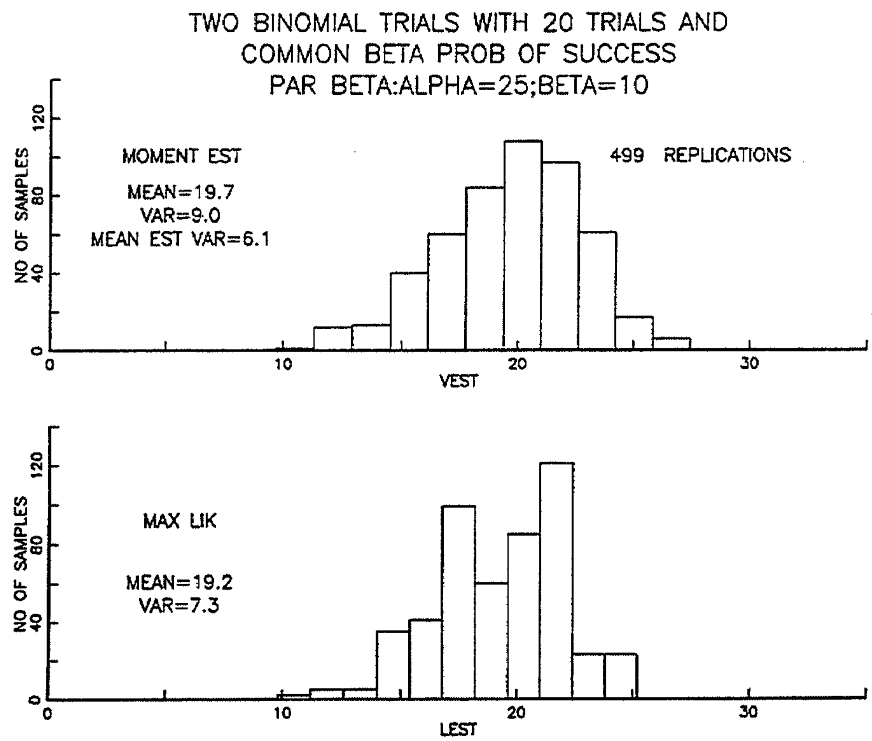


Figure 21

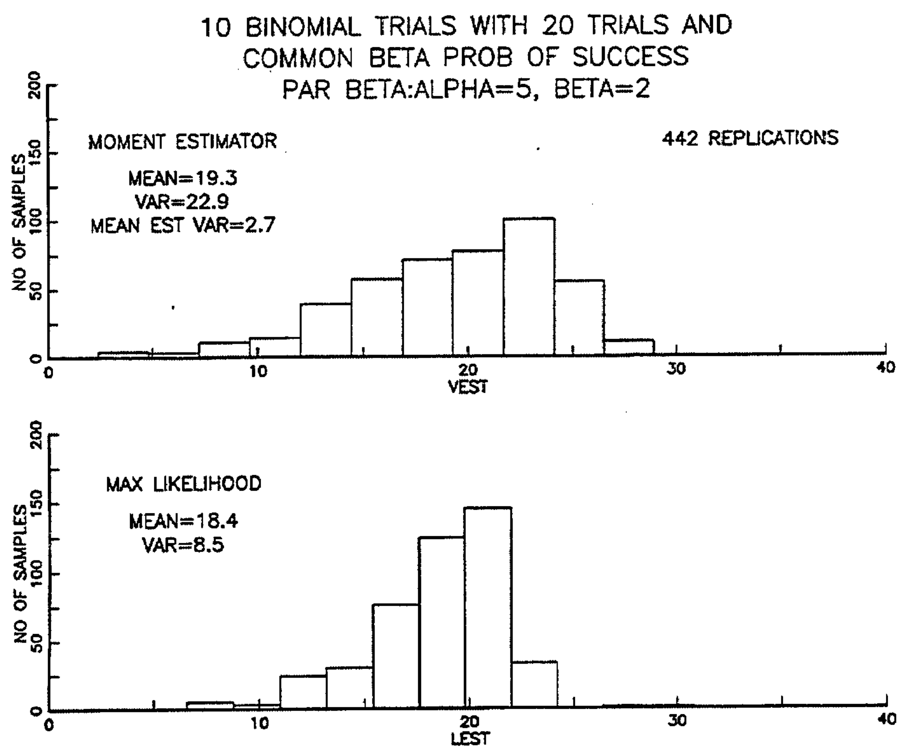


Figure 22

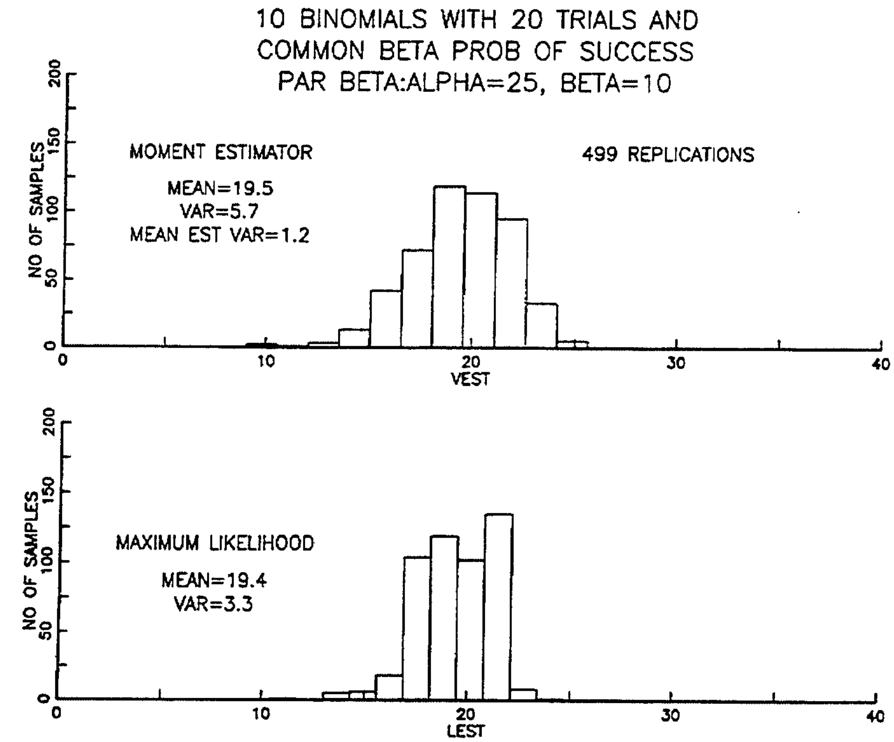


Figure 23

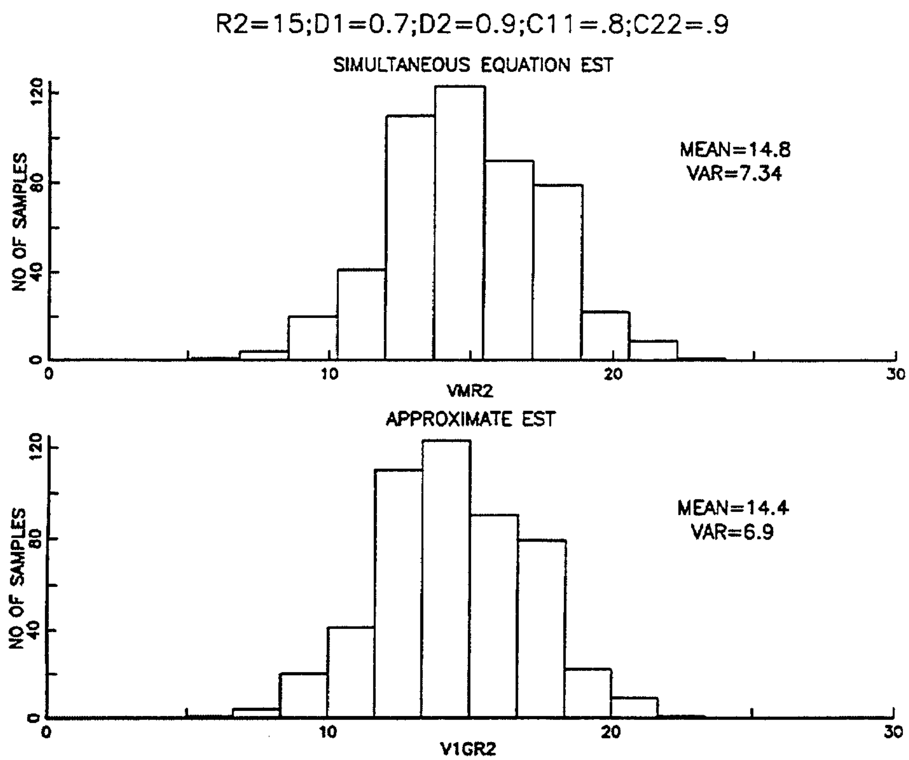


Figure 24

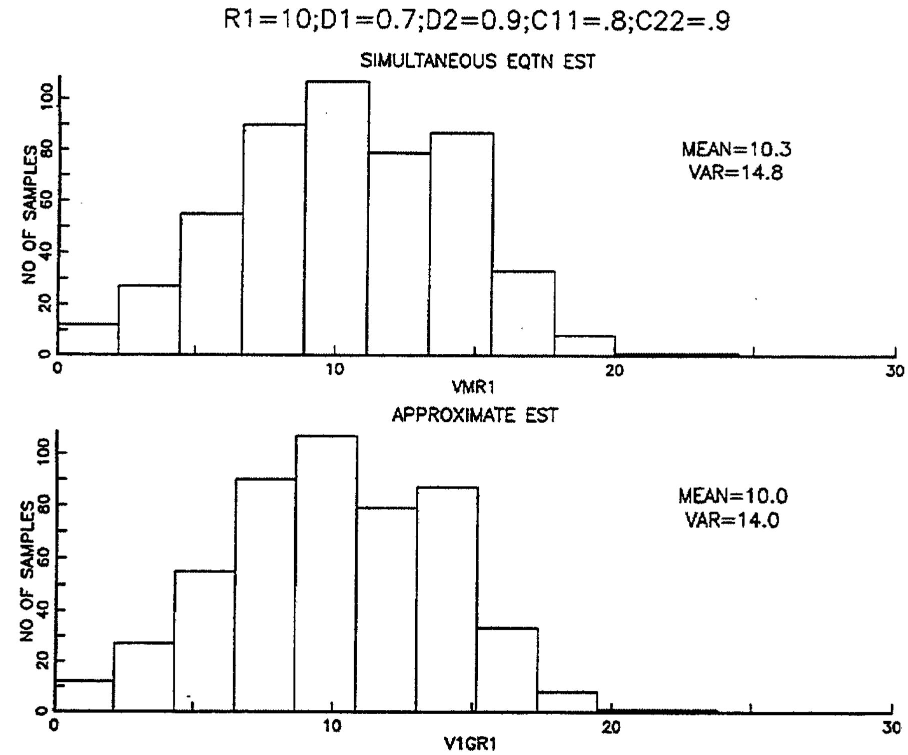


Figure 25

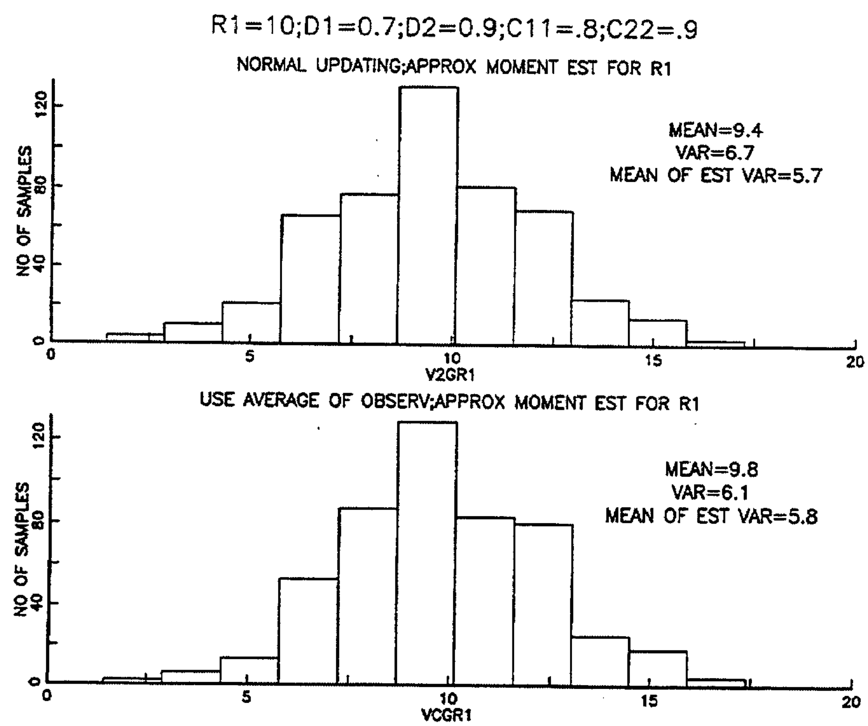


Figure 26

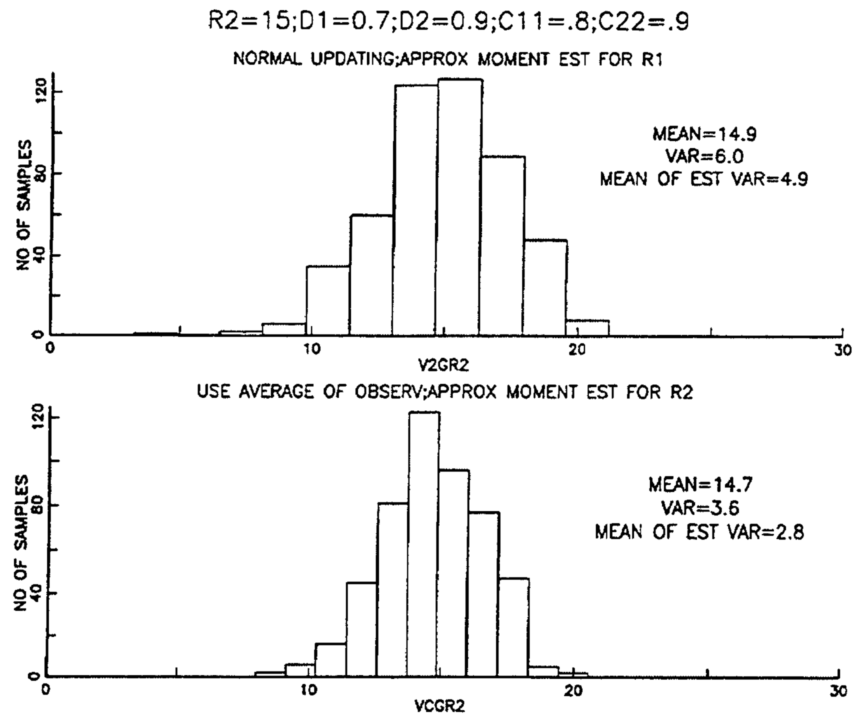

R2=15;D1=0.7;D2=0.9;C11=.8;C22=.9
NORMAL UPDATING;APPROX MOMENT EST FOR R1
NO OF SAMPLES
MEAN=14.9
VAR=6.0
MEAN OF EST VAR=4.9
V2GR2
USE AVERAGE OF OBSERV;APPROX MOMENT EST FOR R2
NO OF SAMPLES
MEAN=14.7
VAR=3.6
MEAN OF EST VAR=2.8
VCGR2

Figure 27

REFERENCES

Basseville, M., "Detecting changes in signals and systems – a survey," Automatica, 24 (1988) pp. 309-326.

Buckland, S.T., Anderson, D.R., Burnham, K.P., and Laake, J.L., *Distance Sampling. Estimating abundance of biological populations*. Chapman and Hall, London, 1993.

Feller, W., *An Introduction to Probability Theory and its Applications*, Vol I, Third Edition, 1968.

Gaver, D.P., "Random hazard in reliability problems," *Technometrics*, **5** (1963) pp. 211-226.

Hall, P., "On the erratic behavior of estimators of N in the binomial N, p distribution," *J. Amer. Statist. Assoc.*, **89** (1994) pp. 344-352.

Helmbold, R.L., "Decision in battle: breakpoint hypotheses and engagement termination data," RAND Technical Report R-772-PR, Rand Corporation, Santa Monica, CA 90406, June 1971.

Hughes, W.P. Jr., "Uncertainty in Combat," *Military Operations Research*, Summer (1994) pp. 45-57.

Lai, T.L., "Sequential changepoint detection in quality control and dynamical systems," *J. R. Statist. Soc.* B, **57** (1995) pp. 613-658.

McCullagh, P. and Nelder, J.A., *Generalized Linear Models*, Chapman and Hall, New York, 1983.

Tierney, L. and Kadane, J.B., "Accurate approximation for posterior moments and marginal densities," *J. Amer. Statis. Assoc.*, **81**, (1986) pp. 82-86.

von Clausewitz, C., *On War*, edited and translated by Michael Howard and Peter Paret, Princeton University Press, Princeton, NJ, 1976.

Youngren, M.A., The Joint Warfare Analysis Experimental Prototype (JWAEP), User Documentation (Draft), unpublished, 1996.

APPENDIX A

Node Occupancy and Its Perception, and COA Inference

Describe the *occupancy* of a node, denoted by n, by the numbers of units of different types (e.g., heavy armor brigades, light armor brigades, etc.); $U_i(n; t)$ is the number of such units of type i ($i = 0, 1, 2, \ldots, I$) at time t at node n; $U_0(n; t)$ can designate the empty node if necessary.

Within a unit of type i, there may be several asset types. Agree that distinguishable assets occur in J classes, and that a unit of type i has a mean number of assets of type j ($j = 1, 2, \ldots, J$) equal to $\alpha_{ij}(n,t)$, and a variance $\sigma_{ij}^2(a;n,t)$. Furthermore, adopt the provisional model that the actual number of assets of type j possessed by a particular randomly selected unit is a random variable, $A(i, j, n, t)$ with distribution function $F_{ij}(x; n, t)$. When convenient, and for illustration, we take $A_k(i, j; n, t)$, $k = 1, 2, \ldots, U_i$, i.e., the numbers of assets of type j owned by the U_i copies on Node n to be independent and normally/Gaussian distributed. The time-dependent parameters α_{ij} and σ_{ij}^2 can reflect the fact that a campaign has been in progress for some time and attrition has occurred and is subject to change. Let

$$\overline{A}(i,j,n,t) = \sum_{\ell=1}^{U_i} A_\ell(i,j,n,t) \tag{A.1}$$

denote the total number of assets of type j possessed by units of type i at the n^{th} node at time t.

It is convenient to refer to the vector of distributions of typical asset-type numbers for a particular unit type as the *signature* (or *asset signature*) of the unit type. Note that signatures of different individual units of the same type will inevitably differ if their various asset counts differ, as could well happen. Let

$$\overline{A}(j,n,t)=\sum_{i=1}^{I}\sum_{\ell=1}^{U_i}A_\ell(i,j,n,t), \tag{A.2}$$

the total number of assets of type j at the node at time t.

Finally, let $\boldsymbol{S}_j(s, n; t)$ denote the total count of assets of type j by sensors of type s ($s = 1, 2, \ldots, S$) at node n at time t. $\boldsymbol{S}_j$ represents the quantitative *perception* of the opponent's type j asset level. It is convenient, although possibly optimistic, to assume initially that

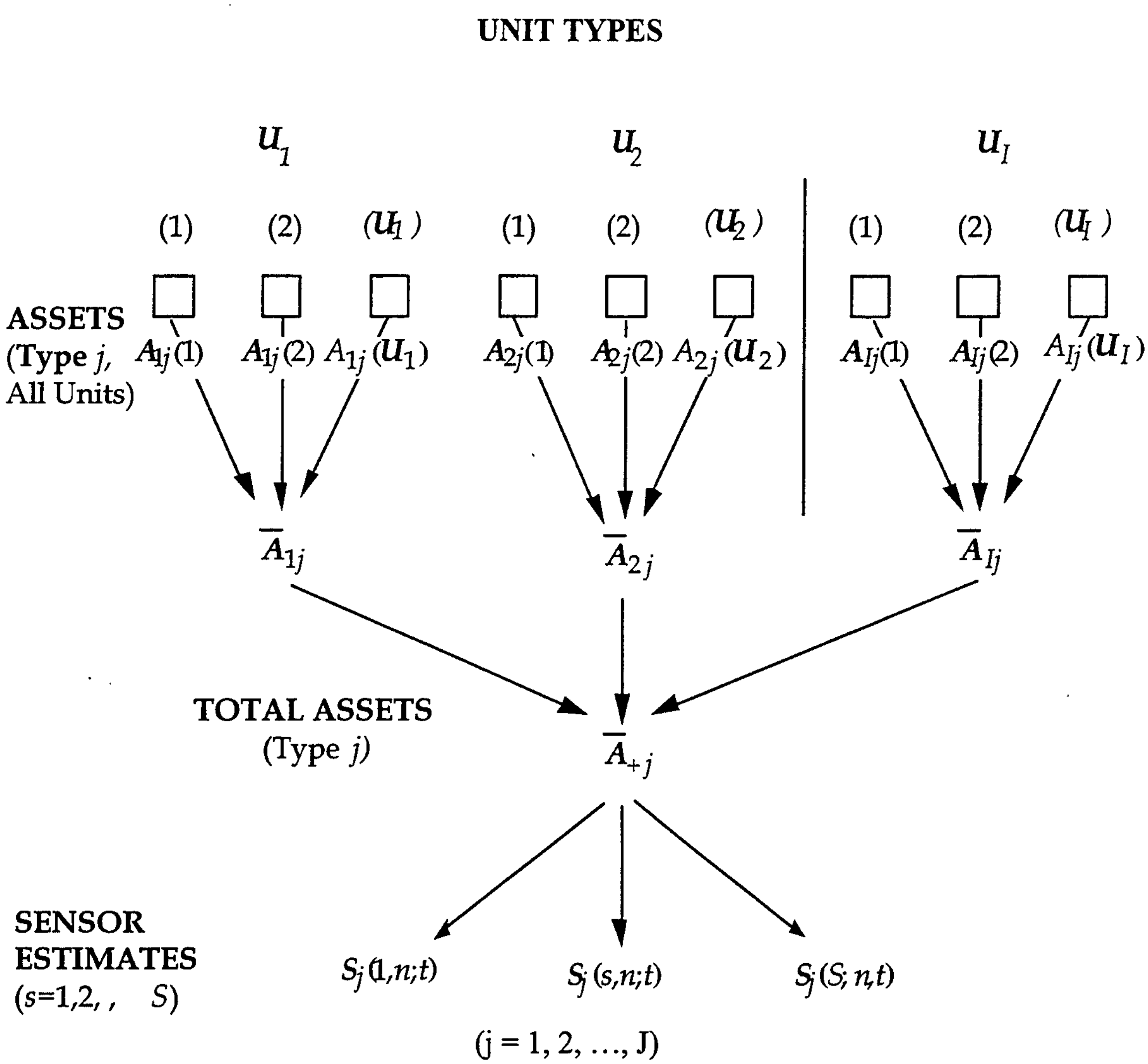
$$E\left[\boldsymbol{S}_j(s,n;t)\middle|\overline{A}(j,n,t)\right]=\overline{A}(j,n,t) \tag{A.3}$$

with a variance that can depend upon sensor type s, node identity, n, and the time (especially if at night). Other influences on the ability of a sensor to assess the presence and strength of an opponent's assets are the latter's activity: radiation using radars, or communication attempts are two such clue-providers.

Towards Probabilistic Node-Asset Assessment, Given Perception

In an idealistic and theoretical sense it would be desirable for a protagonist in a theater-level campaign to possess the joint probability distribution of the types of units occupying all relevant nodes, along with their asset positions, at every time t. This distribution could be influenced by the protagonist's disposition of its sensor assets and own force maneuver as well as by the opponent's maneuver, presumably influenced by *its* own sensor asset disposition. Of course the raw sensor data can, and almost certainly will, be affected by deceptive tactics employed by the other side. Consequently fusion of the data and subsequent actions must attempt to wisely account for such a likelihood when situation assessments are made and future actions planned.

The development of an automated systemic model of a theater campaign, or even a modest fragment thereof, must realistically stop short of even a completely integrated joint probabilistic assessment of nodal state for a particular protagonist. Instead, we will endeavor to develop plausible *marginal* probabilistic descriptions of individual nodal states.



The Unit-Types, Assets, Sensors Hierarchy.

A Univariate Model

We start with a univariate model, both for assets and sensors. For simplicity we drop explicit reference to n and t. Suppose that the node is occupied by (possibly multiple) units of one type, i, and a sensor of type s estimates the number of assets of type j at that node. We assume the following model: First, sensor assessment is randomly distributed around the actual asset number:

$$P\left\{S_j(s;t) \in dx \middle| \overline{A}(j) = a\right\} = \frac{1}{\sqrt{2\pi}\tau_{s,j}} \exp\left\{-\tfrac{1}{2}(x-a)^2 / \tau_{s,j}^2\right\} dx \qquad \text{(A.4)}$$

Second, asset number is randomly distributed around a known value, $\alpha_{ij}(n, t)$, where the latter is derivable from the Table of Organization and Equipment (TOE), plus any information available concerning unit campaign experience (the unit may have been attrited). *Assuming* that actual j^{th}-asset values are independently distributed around the TOE value, then

$$P\left\{\overline{A}(j) \in da \middle| U_i = u, U_k = 0, k \neq i\right\} = \frac{1}{\sqrt{2\pi}\sqrt{u\sigma_{ij}(a)}} \exp\left\{-\tfrac{1}{2}(a-u\alpha_{ij})^2 / u\sigma_{ij}^2(a)\right\} da \quad \text{(A.5)}$$

It can easily be shown that

$$P\left\{S_j(s,t) \in dx \middle| U_i = u, U_k = 0, k \neq i\right\} =$$

$$\frac{1}{\sqrt{2\pi}\sqrt{u\sigma_{ij}^2(a) + \tau_{sj}^2}} \exp\left\{-\frac{1}{2}\frac{1}{u\sigma_{ij}^2(a) + \tau_{sj}^2}(x-u\alpha_{ij})^2\right\} dx \qquad \text{(A.6)}$$

Further, by Bayes' formula,

$$P\left\{\overline{A}(j) \in da \middle| S_j(s;t) = x, U_i = u, U_k = 0, k \neq i\right\}$$

$$= \frac{1}{\sqrt{2\pi}v_j(s,i,u,x)} \exp\left\{-\frac{1}{2}(a - m_j(s,i,u,x))^2 / v_j^2(s,i,u,x)\right\} da \qquad \text{(A.7)}$$

where

$$v_j^2(s,i,u,x) = \frac{u\sigma_{ij}^2(a)\tau_{sj}^2}{u\sigma_{ij}^2(a)+\tau_{sj}^2} \tag{A.8}$$

$$m_j(s,i,u,x) = \frac{u\sigma_{ij}^2(a)}{u\sigma_{ij}^2(a)+\tau_{sj}^2}x + \frac{\tau_{sj}^2}{u\sigma_{ij}^2(a)+\tau_{sj}^2}u\alpha_{ij}\,. \tag{A.9}$$

It next becomes necessary to account for the unknown number, u, of i-type units assumed to be present in order to assess the asset count distribution. To do so, let

$$\Pi(i,u) = P\{U_i = u, U_k = 0, k \neq i\} \tag{A.10}$$

the prior probability that there are u units of just one type, i, at the node. Recall that we are assuming there is only one type of unit at the node; thus,

$$\sum_{i=1}^{I}\sum_{u}\Pi(i,u) = 1 \tag{A.11}$$

We must compute

$$\begin{aligned}\Pi(i,u;x) &= P\{U_i = u, U_k = 0, k \neq i | S_j(s;t) = x\}\\ &= \int P\{U_i = u, U_k = 0, k \neq i, A_j \in da | S_j(s;t) = x\}\\ &= c\Pi(i,u)\frac{1}{\sqrt{2\pi}\sqrt{u\sigma_{ij}^2(a)+\tau_{sj}^2}}\exp\left\{-\tfrac{1}{2}\frac{1}{u\sigma_{ij}^2(a)+\tau_{sj}^2}(x-u\alpha_{ij})^2\right\}\end{aligned} \tag{A.12}$$

where the constant c is determined by the normalization condition.

$$\sum_{i}\sum_{u}\Pi(i,u;x) = 1 \tag{A.13}$$

Finally, then, the distribution of assets of type j is a probability (or convex) combination of normals:

$$\begin{aligned}&P\{\overline{A}(j) \in da | S_j(s,t) = x\} =\\ &\sum_{i,u}\left[\frac{1}{\sqrt{2\pi}\upsilon_j(s,i,u,x)}\exp\left\{-\frac{1}{2}(a-m_j(s,i,u,x))^2 / \upsilon_j^2(s,i,u,x)\right\}\right]\Pi(i,u;x)da\end{aligned} \tag{A.14}$$

and the joint distribution

$$P\left\{U_i = u, U_k = 0, k \neq i, \bar{A}(j) \in da \middle| S_j(s;t) = x\right\}$$

$$= \Pi(i,u;x)\frac{1}{\sqrt{2\pi}\upsilon_j(s,i,u,x)}\exp\left\{-\frac{1}{2}\left(a - m_j(s,i,u,x)\right)^2 / \upsilon_j^2(s,i,u,x)\right\}da. \tag{A.15}$$

This latter joint distribution is of the same form as the joint distribution prior to the sensor estimate. Thus similar calculations can be used to update the distribution as new sensor observations come in. To be specific let

$$\Pi(i,u;x(t),t) = P\left\{U_i = u, U_k = 0, k \neq i \middle| S_j(s(w),t(w)) = x(w), t(w) \leq t\right\} \tag{A.16}$$

be the conditional distribution of the type of unit and number of that type of unit present given all the sensor observations that have occurred before time t. Let $m_j(t) \equiv m_j(s,i,u,x;t)$ (respectively $\upsilon_j^2(t) \equiv \upsilon_j^2(s,i,u,x,t)$) be the conditional expected value (respectively variance) of the number of asset j given there are u units of type i present at the node and all of the sensor observations before time t. Suppose a new sensor observation $x(t+1)$ from a sensor of type s occurs at $t+1$. The posterior distribution Π is updated as follows

$$\Pi(i,u;x(t+1),t+1)$$

$$= C\Pi(i,u;x(t),t)\xi\left(x(t+1); m_j(s,i,u,x(t);t), \left(\upsilon_j^2(s,i,u,x(t);t) + \tau_{sj}^2\right)^{0.5}\right) \tag{A.17}$$

where $\xi(x; m, \upsilon)$ is the normal density function with mean m and standard deviation υ evaluated at x and C is a normalizing constant. Further the conditional moments of the number of asset j given the sensor observations and the type and number of unit present are updated as follows

$$m_j(t+1) = \frac{\upsilon_j^2(t)}{\upsilon_j^2(t)+\tau_{sj}^2}x(t+1)+\frac{\tau_{sj}^2}{\upsilon_j^2(t)+\tau_{sj}^2}m_j(t) \tag{A.18}$$

and

$$\upsilon_j^2(t+1)=\frac{\upsilon_j^2(t)\tau_{sj}^2}{\upsilon_j^2(t)+\tau_{sj}^2}$$

An Independent Multivariate Model

Once again assume that the node is occupied by (possibly multiple) units of one type and a sensor of type s estimates the number of assets of various types (say 2) at the node. Consider the following model. First a sensor assesses the numbers of different assets independently and each sensor asset assessment is randomly distributed around the actual asset number.

$$P\left\{S_1(s;t)\in dx_1, S_2(s;t)\in dx_2 \middle| \overline{A}(1)=a_1, \overline{A}(2)=a_2\right\} = \prod_{j=1}^{2}\frac{1}{\sqrt{2\pi}\tau_{s,j}}\exp\left\{-\tfrac{1}{2}\left(x_j-a_j\right)^2/\tau_{s,j}^2\right\}dx_j \tag{A.19}$$

Second, given the type of unit and the number of units present, the counts of the different types of assets are independently distributed around the TOE values.

$$P\left\{\overline{A}(1)\in da_1, \overline{A}(2)\in da_2 \middle| U_i=u, U_k=0, k\neq i\right\} = \prod_{j=1}^{2}\frac{1}{\sqrt{2\pi}\sqrt{u}\sigma_{ij}(a)}\exp\left\{-\tfrac{1}{2}\left(a_j-u\alpha_{ij}\right)^2/u\sigma_{ij}^2(a)\right\}da_j \tag{A.20}$$

One can calculate,

$$\begin{aligned}
&P\left\{S_j(s;t)\in dx_j; j=1,2 \middle| U_i=u, U_k=0, k\neq i\right\} \\
&=\int_{a_1}\int_{a_2}P\left\{S_j(s;t)\in dx_j; \overline{A}(j)\in da_j; j=1,2 \middle| U_i=u, U_k=0, k\neq i\right\} \qquad \text{(A.21)} \\
&=\prod_{j=1}^{2}\frac{1}{\sqrt{2\pi}\sqrt{u\sigma_{ij}^2(a)+\tau_{sj}^2}}\exp\left\{-\tfrac{1}{2}\frac{1}{u\sigma_{ij}^2(a)+\tau_{sj}^2}\left(x_j-u\alpha_{ij}\right)^2\right\}dx_j
\end{aligned}$$

Thus, if

$$\Pi(i,u) = P\{U_i = u, U_k = 0, k \neq i\} \tag{A.22}$$

as before, then

$$\begin{aligned} &P\{U_i = u, U_k = 0, k \neq i | S_1(s;t) = x_1, S_2(s;t) = x_2\} \\ &= C\Pi(i,u)\prod_{j=1}^{2} \frac{1}{\sqrt{2\pi}\sqrt{u\sigma_{ij}^2(a) + \tau_{sj}^2}} \exp\{-\tfrac{1}{2}\frac{1}{u\sigma_{ij}^2(a) + \tau_{sj}^2}(x_j - u\alpha_{ij})^2\} \\ &= \Pi(i,u;x_1,x_2). \end{aligned} \tag{A.23}$$

Further

$$\begin{aligned} &P\{\overline{A}(j) \in da_j; j = 1,2 | U_i = u, U_k = 0, k \neq i, S_j(s,t) = x_j; j = 1,2\} \\ &= \prod_{j=1}^{2} \frac{1}{\sqrt{2\pi} v_j(s,i,u,x_j)} \exp\{-\tfrac{1}{2}(a_j - m_j(s,i,u,x_j))^2 / v_j(s,i,u,x_j)^2\} da_j \end{aligned} \tag{A.24}$$

where

$$\begin{aligned} v_j^2(s,i,u,x_j) &= \frac{u\sigma_{ij}^2(a)\tau_{sj}^2}{u\sigma_{ij}^2(a) + \tau_{sj}^2} \\ m_j(s,i,u,x_j) &= \frac{u\sigma_{ij}^2(a)}{u\sigma_{ij}^2(a) + \tau_{sj}^2} x_j + \frac{\tau_{sj}^2}{u\sigma_{ij}^2(a) + \tau_{sj}^2} u\alpha_{ij} \end{aligned} \tag{A.25}$$

Letting

$$\xi(x;m,v) = \frac{1}{\sqrt{2\pi} v} \exp\left\{-\frac{1}{2}(x-m)^2 / v^2\right\}, \tag{A.26}$$

the posterior distribution

$$P\left\{\overline{A}(j) \in da_j; j=1,2, U_i=u, U_k=0, k \neq i \middle| S_j(s;t)=x_j; j=1,2\right\}$$

$$= C\Pi(i,u)\prod_{j=1}^{2}\xi\left(a_j; m_j(s,i,u,x_j), v_j(s,i,u,x_j)\right)\xi\left(x_j; u\alpha_{ij}, \left(u\sigma_{ij}^2+\tau_{sj}^2\right)^{0.5}\right)da_j \quad \text{(A.27)}$$

$$= \Pi(i,u;x_1,x_2)\prod_{j=1}^{2}\xi\left(a_j; m_j(s,i,u,x_j), v_j(s,i,u,x_j)\right)da_j$$

where

$$v_j^2(s,i,u,x_j) = \frac{u\sigma_{ij}^2(a)\tau_{sj}^2}{u\sigma_{ij}^2(a)+\tau_{sj}^2}$$

$$m_j(s,i,u,x_j) = \frac{u\sigma_{ij}^2(a)}{u\sigma_{ij}^2(a)+\tau_{sj}^2}x_j + \frac{\tau_{sj}^2}{u\sigma^2(a)+\tau_{sj}^2}u\alpha_{ij} \quad \text{(A.28)}$$

Recall that,

$$P\{U_i=u, U_k=0, k\neq i, A_j \in da_j; j=1,2\}$$

$$= \Pi(i,u)\prod_{j=1}^{2}\xi\left(a_j; u\alpha_{ij}, u\sigma_{ij}^2(a)\right)da_j \quad \text{(A.29)}$$

and thus the posterior distribution is of the same form as the prior. Hence, one can use the above posterior distribution as the new prior and compute a new posterior distribution when new sensor information becomes available. To be specific, let

$$\Pi(i,u;x(t),t)$$

$$= P\{U_i=u, U_k=0, k\neq i \mid S_j(s(w),t(w)) = x(w), t(w)\leq t, j=1,2\} \quad \text{(A.30)}$$

be the conditional distribution of the type of unit and the number of that type of unit present at the node given all the sensor observations that have occurred before time t. Let $m_j(t) = m_j(s,i,u,\mathbf{x};t)$ (respectively $v_j^2(t) = v_j^2(s,i,u,\mathbf{x},t)$) be the conditional expected value (respectively variance) of the number of asset j given there are u units of type i present at the node and all of the sensor observations before time t. Suppose a new sensor observation $\{x_j(t+1)\}$ from a sensor of type s occurs at $t+1$. The posterior distribution Π is updated as follows

$$\Pi(i,u;\boldsymbol{x}(t+1),t+1) = C\Pi(i,u;\boldsymbol{x}(t),t)\prod_j \xi\left(x_j(t+1), m_j(s,i,u,x;t), \left[v_j^2(t)+\tau_{sj}^2\right]^{0.5}\right) \tag{A.31}$$

where the product is over those types of assets for which there is an observation; $\xi(x; m, \upsilon)$ is the normal density function with mean m and standard deviation υ evaluated at x and C is a normalizing constant. Further the conditional moments of the number of asset j given the sensor observations and the type and number of unit present are updated as follows

$$m_j(t+1) = \frac{v_j^2(t)}{v_j^2(t)+\tau_{sj}^2}x(t+1) + \frac{\tau_{sj}^2}{v_j^2(t)+\tau_{sj}^2}m_j(t)$$

and (A.32)

$$v_j^2(t+1) = \frac{v_j^2(t)\tau_{sj}^2}{v_j^2(t)+\tau_{sj}^2}$$

Example

In this example there are 2 types of units: I and II. There are 3 asset types. Suppose unit I has 0 assets of type 1, 100 assets of type 2 and 30 assets of type 3. Suppose unit II has 50 assets of type 1, 150 assets of type 2 and 50 assets of type 3.

Suppose there are 2 sensor types. Sensor type 1 can estimate the count of assets of type 1 and 2, but not 3. Sensor type 3 can estimate the count of assets of type 3 but not 1 and 2.

Assume the following prior distribution.

$$P\left\{\overline{A}(j) \in da_j \middle| U_i = u; U_k = 0, k \neq i\right\} =$$

$$\prod_{j=1}^{3} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{u\sigma_{ij}(a)}} \exp\left\{-\tfrac{1}{2}\left(a_j - u\alpha_{ij}\right)^2 / u\sigma_{ij}^2(a)\right\} da_j \tag{A.33}$$

$$P\{U_i = u, U_k = 0, k \neq i\} = \frac{1}{4} \quad \text{for} \quad i = 1,2; u = 1,2$$

where

$$\alpha_{1j} = \begin{cases} 0 & \text{if } j=1 \\ 100 & \text{if } j=2 \\ 30 & \text{if } j=3 \end{cases} \qquad \sigma_{1j} = \begin{cases} 0.5 & \text{if } j=1 \\ 25 & \text{if } j=2 \\ 5 & \text{if } j=3 \end{cases}$$

$$\alpha_{2j} = \begin{cases} 75 & \text{if } j=1 \\ 150 & \text{if } j=2 \\ 50 & \text{if } j=3 \end{cases} \qquad \sigma_{2j} = \begin{cases} 10 & \text{if } j=1 \\ 25 & \text{if } j=2 \\ 5 & \text{if } j=3 \end{cases} \tag{A.34}$$

that is, the number of assets of each type are independently normally distributed and the prior distribution of the unit type and number of units occupying the node assigns probability 0.25 to each of the events: there is one (respectively two) units of type I; there is one (respectively two) units of type II.

Assume the distribution of the sensor estimate of the count of assets at the node is

$$P\left\{S_j(1,n;t) \in dx_j; j = 1,2 \middle| \overline{A}(j) = a_j; j = 1,2\right\}$$

$$= \prod_{j=1}^{2} \frac{1}{\sqrt{2\pi}\tau_{1,j}} \exp\left\{-\frac{1}{2}\left(x_j - a_j\right)^2 / \tau_{1j}^2\right\} dx_j \tag{A.35}$$

where

$$\tau_{1j} = \begin{cases} 25 & \text{if } j=1 \\ 35 & \text{if } j=2 \end{cases}$$

$$P\left\{S_3(2,n,t) \in dx_3 \middle| A_3 = a_3\right\} = \frac{1}{\sqrt{2\pi}\tau_{23}} \exp\left\{-\frac{1}{2}\left(x_3 - a_3\right) / \tau_{23}^2\right\} \tag{A.36}$$

where $\tau_{23} = 10$.

The results of the updating procedure appears in Table A1. Displayed are the time of the sensor report, the type of sensor making the report, the estimated number of assets of each type made by the sensor, the posterior distribution of the type of unit and number of units present, and the conditional posterior means and standard deviations of the number of assets of each type present given the type of unit and the number of units present.

For time 0, the prior distribution of the unit type and number of units present $\Pi(i, u) = 1/4$ is displayed. Also displayed are the conditional moments of the prior normal distribution of numbers of assets of each type given the unit type and number of units present. At time 1 there is a report from a sensor of type 1 that there are 10 assets of type 1 and 80 assets of type 2. This information updates the probability that there is 1 unit of type I at the node to 0.94. At time 2 there is a report from a sensor of type 2 that there are 35 assets of type 3. This information updates the probability that there is 1 unit of type I at the node to 0.99. At time 3 there is a report from a sensor of type 1 that there are 60 assets of type 1 and 140 assets of type 2. This information decreases the probability that there is one unit of type I at the node to 0.81 and increases the probability that there is one unit of type 2 at the node to 0.18. At time 4 there is a report from a sensor of type 1 that there are 90 assets of type 1 and 200 assets of type 2. This information increases the probability that there is one unit of type II at the node to 1. At time 5 there is a report from a sensor of type 2 that there are 70 assets of type 3. This information decreases the probability that there is one unit of type II at the node to 0.13 and increases the probability that there are 2 units of type II at the node to 0.87.

The sensor observations were chosen so that the observations at times 1 and 2 might come from one unit of type I. The sensor observations after time 2 may come from another situation. Note that the procedure does respond to the change.

To further investigate the responsiveness of the procedure to change, a simulation was conducted. The ground truth for the simulation is as follows. For the first 5 time periods there is 1 unit of type I at the node. The (true) count of each asset type is drawn from the prior normal distribution for one unit of type I and rounded down to the closest integer; if the simulated asset count is negative it is set equal to zero. The (true) count of asset 1 (respectively 2, 3) is 0 (respectively 69, 38) for the first five time periods. For the remainder of the time periods, there are 2 units of type II at the node. Once again the count of assets of each type is drawn from the prior normal distribution for two units of type II and rounded down to the closest integer. The resulting count of asset 1 (respectively 2, 3) is 149 (respectively 290, 108). The standard deviation of sensor 1 on asset 1 (respectively 2) is 25 (respectively 35). The standard deviation of sensor 2 on asset 3 is 10.

For each time period, a random number is drawn to determine the type of sensor making the observation; (each sensor type is equally likely). Given the sensor type, the sensor observation is drawn from a normal distribution with mean the true asset number and standard deviation for the asset type and sensor type; the resulting number is set equal to zero if it is negative.

The results of the simulation appear in Table A2. Once again the table lists the times of the observations, the type of sensor making the observation, the sensor's estimated number of assets, the posterior distribution of the unit type and number of units after each observation, and the conditional posterior moments. The row for time 0 displays the prior distribution.

Recall that there is one unit of type I at the node for times 1–5. The posterior distributions quickly reflect this ground truth; after the first observation the posterior probability is 0.89 that there is one unit of type I; after the second observation, this posterior probability is increased to 0.99; after the third observation to 1.0.

Recall that for times 6 on, there are two units of type II at the node. The observation at time 6 causes the posterior distribution of there being 1 unit of type I to decrease to 0.05 and increases the probability of there being 2 units of type I to 0.95. Thus, the updating procedure does reflect the fact that a change has occurred but does not correctly identify the type and number of units. It is not until after the observation at time 12, that the updating procedure puts a sizable posterior probability on the event that there are 2 units of type II. This sluggish behavior is probability due to that fact that after the observation at time 5, the posterior probability that there are two units of type II at the node is (a very unlikely) 10^{-26}. Subsequent observations raise the probability to 10^{-5} after the 11^{th} observation. Thus, it appears that the Bayesian methodology must overcome its belief after time 5 that it is extremely unlikely that there are 2 units of type II at the node.

Some Generalizations

Let $X(n;t) = (U_1(n;t), U_2(n;t) \ldots, U_I(n;t))$ be a vector describing the number of units of each type at node n at time t; $U_i(n,t)$ is the number of units of type i at node n at time t. If the node is unoccupied at time t, $X(n,t) = (0,0,\ldots,0)$.

The total number of assets of type j at the node at time t is

$$\overline{A}(j,n,t) = \sum_{i=1}^{I} \sum_{k=1}^{U_i(n,t)} A_k(i,j,n,t).$$

Assume $\{A_k(i,j,n,t)\}$ are independent normal random variables with mean α_{ij} and variance $\sigma_{ij}^2(a)$. If $U_i(n,t) > 0$ for some i then $\overline{A}(j,n,t)$ is normal with mean

$$m_j(u_1 \ldots, u_I; n; 0) = \sum_{i=1}^{I} \sum_{k=1}^{u_i} \alpha_{ij}$$

and variance

$$\upsilon_j^2(u_1 \ldots, u_I; n; 0) = \sum_{i=1}^{I} \sum_{k=1}^{u_i} \sigma_{ij}^2(a)$$

If the node is empty we assume $\overline{A}(j,n,t)$ is normal with mean 0 and small standard deviation, e.g., 0.25.

Let $\pi(u_1, u_2, ..., u_I; 0)$ denote the prior possibility that there are u_i units of type i present at the node.

Assume that a sensor of type s is able to estimate the number of assets of type j for $j \in J(s)$ where $J(s)$ is a subset of the assets types. For each sensor type s

$$P\left\{S_j(s;n;t) \in dx_j; j \in J(s) \middle| \overline{A}(j,n,t) = a_j; j \in J(s)\right\}$$

$$= \prod_{j \in J(s)} \frac{1}{\sqrt{2\pi}\tau_{s,j,n,t}} \exp\left\{-\frac{1}{2}\left((x_j - a_j)/\tau_{s,j,n,t}\right)^2\right\};$$

that is the sensor observations are independent normally distributed about the true number of asset j.

Once again a recursive procedure for updating the distribution of the number and types of units at the node can be obtained; as well as a procedure for updating the conditional mean and variance of the number of asset j at the node given the number and types of units. The development is similar to that of the previous sections. To be specific let

$$\Pi((u_1, \ldots, u_I); x(t), t)$$

$$= P\left\{U_1 = u_1, \ldots, U_I = u_I \middle| S_j(s(w), t(w)) = x(w), j \in J(s), t(w) \le t\right\}$$

be the conditional distribution of the type of unit and the number of that type of unit present at the node given all the sensor observations that have occurred before time t. Let $m_j((u_1, ..., u_I); t)$ (respectively $\upsilon_j^2(u_1, \ldots, u_I); t$) be the conditional expected value (respectively variance) of the number of assets of type j given $U_1 = u_1, ..., U_I = u_I$ and all the sensor observations before time t. Suppose a new sensor observation $\{x_j(t+1); j \in J(s)\}$ becomes available. The updated probability of unit type and number is computed as

$$\Pi\left(\left(u_1,\ldots,u_I\right);\boldsymbol{x}(t),\boldsymbol{x}(t+1);t+1\right)$$
$$= C\Pi\left(\left(u_1,\ldots,u_I\right);\boldsymbol{x}(t);t\right)\prod_{j\in J(s)}\xi\left(x_j(t+1);m_j(t),\left(\upsilon_j^2(t)+\tau_{sj}^2\right)^{0.5}\right)$$

where C is a normalizing constant. The conditional moments are updated as follows.

$$m_j\left(\left(u_1,\ldots,u_I\right);t+1\right)$$
$$=\frac{\upsilon_j^2\left(\left(u_1,\ldots,u_I\right);t\right)}{\upsilon_j^2\left(\left(u_1,\ldots,u_I\right);t\right)+\tau_{sj}^2}x_j(t+1)+\frac{\tau_{sj}^2}{\upsilon_j^2\left(\left(u_1,\ldots,u_I\right);t\right)+\tau_{sj}^2}m_j\left(\left(u_1,\ldots,u_I\right);t\right)$$

and

$$\upsilon_j^2\left(\left(u_1,\ldots,u_I\right);t+1\right)=\frac{\upsilon_j^2\left(\left(u_1,\ldots,u_I\right);t\right)\tau_{sj}^2}{\upsilon_j^2\left(\left(u_1,\ldots,u_I\right);t\right)+\tau_{sj}^2}.$$

A Modified Bayesian Approach

The simulation results of Table A2 suggest that the complete Bayesian approach can be rather sluggish in identifying the true number and types of units at a node. Simulation experiments not reported here suggest that in some cases the procedure may never identify the true number and types of units. As mentioned earlier, one cause for this behavior appears to be that the posterior distribution of the number and types of units at the node can attribute essentially 0 probability to particular configurations; this 0 probability can make it difficult for the procedure to adjust to new circumstances.

Another aspect of the Bayesian procedure is that it continually updates the conditional means and variances of the amount of each type of asset for each unit type and number configuration. This continual updating may result in the posterior conditional moments being far away from the original conditional moments; this may negatively influence the ability of the Bayesian procedure to correctly identify the numbers and types of units at a node in a new situation.

Preliminary experiments have suggested that the following may result in an improved ability to correctly identify the number and type of units at a node. At each time t compare the just computed posterior distribution $\pi((u_1, \ldots, u_I);t)$ to the prior distribution $\pi((u_1, \ldots, u_I);t\text{-}1)$. If

$$\max_{(u_1,\ldots,u_I)} \left|\Pi((u_1,\ldots,u_I);t) - \Pi((u_1,\ldots,u_I);t-1)\right| > 0.8,$$

say, declare that a change has occurred. Use an equally likely prior distribution for unit types and numbers when the next sensor observation arrives; also reset the conditional moments to their original values before updating.

An Example

Tables A3 and A4 report the results of a simulation study. The scenario is as follows. One unit of type I has a mean number of asset 1 (respectively assets 2 and 3) of 0 (respectively 100, 30) with standard deviation for the number of asset 1 (respectively assets 2 and 3) of 0.5 (respectively 25 and 5). One unit of type II has a mean number of asset 1 (respectively assets 2 and 3) of 75 (respectively 15 and 50) with standard deviation for the number of asset 1 (respectively assets 2 and 3) of 10 (respectively 25 and 5). A sensor of type 1 can estimate a number of assets of type 1 and type 2 but not type 3; the error standard deviation is 25, (respectively 35) for type 1 (respectively type 2) assets. A sensor of type 2 can estimate numbers of assets of type 3 only with error standard deviation 10.

The possibilities for node occupancy are (0, 0) (empty); (1, 0) (1 unit of type I); (2, 0) (2 units of type I); (0, 1) (1 unit of type II); (0, 2) (2 units of type II), and (1, 1) (1 unit of type I and 1 unit of type II). The prior distribution is that each of these configurations is equally likely.

During the first 5 time periods there is 1 unit of type I at the node; the true number of assets of type 1 (respectively asset type 2 and 3) is 0 (respectively 69 and 39); these numbers are obtained by drawing from the appropriate normal distribution for asset type, making the resulting number 0 if it is negative and rounding down to the next integer value.

During times 6–15 there are 1 unit of type I and 1 unit of type II at the node with 99 combined assets of type 1, 231 combined assets of type 2, and 63 combined assets of type 3; these true numbers of assets are obtained in a similar manner to those of the first 5 time periods.

At each time a sensor type is randomly selected and the sensor observation(s) are drawn from a normal distribution with mean the true number of assets of that type and standard deviation for that asset type for that sensor; the observations are set equal to 0 if they are negative.

Table A3 displays results of using the Bayesian updating procedure. Displayed are the data, the posterior probabilities of each configuration, and the conditional mean number of each asset type given the configuration and the observations. The row at time 0 displays the prior information. The posterior distributions indicate that the procedure uses the sensor observations to correctly classify by time 5 that there is 1 unit of type I at the node. By time 7 the posterior distribution indicates that there is evidence that the occupancy of the node has changed. However, the posterior distribution is never able to assign an appreciable probability to the correct identity of 1 unit of type I and 1 unit of type II.

Table A4 displays the results of using the Bayesian updating procedure with the following adjustment. At each time the prior and posterior distribution are compared. If the total variation is less than 0.8, nothing is done. If it is greater than 0.8, the posterior is replaced with the original prior and the posterior conditional moments are replaced with the original prior conditional moments for the calculations the next time period. In Table A4, the total variation of the prior and posterior exceeded 0.8 at time 11. Thus the results for times 12–15 differ from those of Table A3. Note that in this case the posterior distribution is able to assign a largish probability to the true configuration of 1 unit of type I and 1 unit of type II.

Note that the random assignment of sensor type has assigned a type 2 sensor to make most of the observations. Recall that a type 2 sensor can only estimate numbers of assets of type 3. However, the most obvious feature of a type I unit is that it has no assets of type I. Thus, it is not surprising that the Bayesian procedure is not perfect.

Note also, that even though the occupancy of the node changed at time 6, it was not until time 11 until the total variation between the prior and posterior indicates it; note that time 11 is also the first observation from a sensor of type I after time 5.

Table A1
Results of Updating Perceptions

Time	Sensor Type	Estimated Number of Asset Types			Posterior Distribution (Unit Type, Number)				Conditional Posterior Moments Given (Unit Type, Number) $m_j(s,i,u,x_j)$ $\upsilon_j(s,i,u,x_j)$											
									Unit Type 1						Unit Type 2					
									Nbr = 1 Asset Type			Nbr = 2 Asset Type			Nbr = 1 Asset Type			Nbr = 2 Asset Type		
		1	2	3	(1,1)	(1,2)	(2,1)	(2,2)	1	2	3	1	2	3	1	2	3	1	2	3
0	–	–	–	–	1/4	1/4	1/4	1/4	0	100	30	0	200	60	75	150	50	150	300	100
									(0.50)	(25)	(5)	(0.70)	(35)	(7)	(10)	(25)	(5)	(14)	(35)	(7)
1	1	10	80	–	0.94	0.05	0.01	0	0	93	–	0.01	139	–	66	126	–	116	189	–
									(0.50)	(20)	–	(0.70)	(25)	–	(9)	(20)	–	(12)	(25)	–
2	2	–	–	35	0.98	0.01	0.01	0	–	–	31	–	–	52	–	–	47	–	–	78
									–	–	(4)	–	–	(6)	–	–	(4)	–	–	(6)
3	1	60	140	–	0.81	0.01	0.18	0	0.03	105	–	0.06	140	–	65	130	–	105	172	–
									(0.50)	(18)	–	(0.71)	(20)	–	(9)	(18)	–	(11)	(20)	–
4	1	90	200	–	0	0	1.0	0	0.06	124	–	0.13	155	–	68	144	–	103	179	–
									(0.50)	(16)	–	(0.70)	(17)	–	(8)	(16)	–	(10)	(18)	–
5	2	–	–	70	0	0	0.13	0.87	–	–	38	–	–	56	–	–	51	–	–	76
									–	–	(4)	–	–	(5)	–	–	(4)	–	–	(5)

Table A2. Results of Updating Perceptions

Time	Sensor Type	Estimated Number of Asset Types			Posterior Distribution (Unit Type, Number)				Conditional Posterior Moments Given (Unit Type, Number) $m_j(s,i,u,x_j)$ $\upsilon_j(s,i,u,x_j)$											
									Unit Type 1						Unit Type 2					
									Nbr = 1 Asset Type			Nbr = 2 Asset Type			Nbr = 1 Asset Type			Nbr = 2 Asset Type		
		1	2	3	(1,1)	(1,2)	(2,1)	(2,2)	1	2	3	1	2	3	1	2	3	1	2	3
0	–	–	–	–	1/4	1/4	1/4	1/4	0 (0.25)	100 (25)	30 (5)	0 (0.35)	200 (35)	60 (7)	75 (10)	150 (25)	50 (5)	150 (14)	300 (35)	100 (7)
1	1	0	100	–	0.89	0.10	0.01	0	0 (0.25)	100 (20)	30 (5)	0 (0.35)	150 (25)	60 (7)	65 (9)	133 (20)	50 (5)	114 (12)	199 (25)	100 (7)
2	1	0	43	–	0.99	0.01	0	0	0 (0.25)	86 (18)	30 (5)	0 (0.35)	114 (20)	60 (7)	57 (9)	111 (18)	50 (5)	91 (11)	147 (20)	100 (7)
3	2	–	–	32	1.0	0	0	0	0 (0.25)	86 (18)	30 (4)	0 (0.35)	114 (20)	51 (6)	57 (9)	111 (18)	46 (4)	91 (11)	147 (20)	77 (6)
4	1	0	94	–	1.0	0	0	0	0 (0.25)	87 (16)	30 (4)	0 (0.35)	109 (18)	51 (6)	51 (8)	107 (16)	46 (4)	77 (10)	134 (18)	77 (6)
5	2	–	–	54	0.99	0.01	0	0	0 (0.25)	87 (16)	34 (4)	0 (0.35)	109 (18)	51 (5)	51 (8)	107 (16)	48 (4)	77 (10)	134 (18)	71 (5)
6	2	–	–	93	0.05	0.95	0	0	0 (0.25)	87 (16)	43 (4)	0 (0.35)	109 (18)	60 (4)	51 (18)	107 (16)	54 (4)	77 (10)	134 (18)	76 (4)
7	2	–	–	102	0	1	0	0	0 (0.25)	87 (16)	50 (4)	0 (0.35)	109 (18)	67 (4)	51 (8)	107 (16)	60 (4)	77 (10)	134 (18)	80 (4)
8	1	143	275	–	0	0.66	0.34	0	0.01 (0.25)	119 (14)	50 (4)	0.03 (0.35)	142 (16)	67 (4)	60 (8)	135 (14)	60 (4)	86 (9)	162 (16)	80 (4)
9	1	162	238	–	0	0	1	0	0.03 (0.25)	136 (13)	50 (4)	0.06 (0.35)	158 (14)	67 (4)	69 (7)	150 (13)	60 (4)	95 (9)	175 (14)	80 (4)
10	2	–	–	104	0	0	1	0	0.03 (0.25)	136 (13)	56 (13)	0.06 (0.35)	158 (14)	72 (4)	69 (7)	150 (13)	65 (13)	95 (9)	175 (14)	84 (4)
11	1	162	324	–	0	0	1.0	10^{-5}	0.05 (0.25)	160 (12)	56 (3)	0.09 (0.35)	182 (13)	72 (4)	76 (7)	172 (12)	65 (3)	103 (8)	196 (13)	84 (4)
12	2	–	–	123	0	0	0.77	0.23	0.05 (0.25)	160 (12)	63 (3)	0.09 (0.35)	182 (13)	78 (4)	76 (7)	172 (12)	71 (3)	103 (8)	196 (13)	88 (4)
13	2	–	–	98	0	0	0.15	0.85	0.05 (0.25)	160 (12)	66 (3)	0.09 (0.35)	182 (13)	81 (3)	76 (7)	172 (12)	73 (3)	103 (8)	196 (13)	90 (3)
14	2	–	–	102	0	0	0.01	0.99	0.05 (0.25)	160 (12)	69 (3)	0.09 (0.35)	182 (13)	83 (3)	76 (7)	172 (12)	76 (3)	103 (8)	196 (13)	91 (3)
15	1	152	259	–	0	0	0	1	0.06 (0.25)	171 (12)	69 (3)	0.12 (0.35)	192 (12)	83 (3)	82 (7)	182 (12)	76 (3)	107 (8)	204 (12)	91 (3)

Table A3
Updating Perceptions at a Node; No Adjustment

Time	Sensor Type	Estimated Number of Asset Types			Posterior Distribution (u_1, u_2)						Conditional Posterior Moments Given State of Node $m_j((u_1, u_2), x)$ $\upsilon_j((u_1, u_2), x)$																	
											(1,0) Asset Type			(2,0) Asset Type			(0,1) Asset Type			(0,2) Asset Type			(1,1) Asset Type			(0,0) Asset Type		
		1	2	3	(1,0)	(2,0)	(0,1)	(0,2)	(1,1)	(0,0)	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
0	–	–	–	–	1/6	1/6	1/6	1/6	1/6	1/6	0	100	30	0	200	60	75	150	50	150	300	100	75	250	80	0	0	0
1	1	0	102	–	0.87	0.11	0.01	10^{-10}	10^{-4}	10^{-2}	0	101	30	0	150	60	65	134	50	114	200	100	65	175	80	0	0.01	0
2	2	–	–	48	0.79	0.18	0.03	10^{-14}	10^{-5}	10^{-7}	0	101	34	0	150	56	65	134	50	114	200	82	65	175	69	0	0.01	0.03
3	2	–	–	34	0.95	0.03	0.01	10^{-17}	10^{-7}	10^{-9}	0	101	34	0	150	50	65	134	47	114	200	70	65	175	60	0	0.01	0.05
4	2	–	–	37	0.97	0.02	0.01	10^{-19}	10^{-8}	10^{-12}	0	101	34	0	150	48	65	134	46	114	200	64	65	175	56	0	0.01	0.07
5	1	0	70	–	1.0	10^{-3}	10^{-4}	10^{-25}	10^{-11}	10^{-13}	0	93	34	0	124	48	57	118	46	91	156	64	57	140	56	0	0.01	0.07
6	2	–	–	63	0.94	0.06	10^{-3}	10^{-23}	10^{-9}	10^{-20}	0	93	38	0	124	50	57	118	48	91	156	64	57	140	57	0	0.01	0.11
7	2	–	–	63	0.65	0.34	0.01	10^{-22}	10^{-8}	10^{-27}	0	93	41	0	124	52	57	118	49	91	156	64	57	140	58	0	0.01	0.15
8	2	–	–	48	0.63	0.36	0.01	10^{-23}	10^{-8}	10^{-32}	0	93	41	0	124	52	57	118	49	91	156	62	57	140	57	0	0.01	0.18
9	2	–	–	48	0.59	0.40	0.01	10^{-23}	10^{-8}	10^{-37}	0	93	42	0	124	51	57	118	49	91	156	60	57	140	56	0	0.01	0.21
10	2	–	–	63	0.25	0.74	0.01	10^{-22}	10^{-8}	10^{-45}	0	93	44	0	124	52	57	118	50	91	156	60	56	140	56	0	0.02	0.25
11	1	117	265	–	10^{-4}	0.03	0.97	10^{-18}	10^{-5}	10^{-56}	0.01	128	44	0.02	159	52	63	147	50	96	184	60	63	171	56	0	0.01	0.25
12	2	–	–	83	10^{-5}	0.06	0.94	10^{-17}	10^{-4}	10^{-69}	0.01	128	47	0.02	159	55	63	147	53	96	184	63	63	171	59	0	0.02	0.30
13	2	–	–	58	10^{-5}	0.06	0.94	10^{-17}	10^{-4}	10^{-76}	0.01	128	48	0.02	159	55	63	147	53	96	184	62	63	171	59	0	0.01	0.34
14	2	–	–	62	10^{-5}	0.07	0.93	10^{-17}	10^{-4}	0	0.01	128	49	0.02	159	56	63	147	54	96	184	62	63	171	59	0	0.02	0.38
15	1	106	200	–	10^{-9}	10^{-5}	1.0	10^{-16}	10^{-4}	0	0.02	140	49	0.04	167	56	68	156	54	97	187	62	68	177	59	0	0.01	0.38

Table A4
Updating Perceptions at a Node; With Adjustment

Time	Sensor Type	Estimated Number of Asset Types			Posterior Distribution (u_1, u_2)						Conditional Posterior Moments Given State of Node $m_j((u_1, u_2), x)$ $\upsilon_j((u_1, u_2), x)$																	
											(1,0) Asset Type			(2,0) Asset Type			(0,1) Asset Type			(0,2) Asset Type			(1,1) Asset Type			(0,0) Asset Type		
		1	2	3	(1,0)	(2,0)	(0,1)	(0,2)	(1,1)	(0,0)	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
0	–	–	–	–	1/6	1/6	1/6	1/6	1/6	1/6	0	100	30	0	200	60	75	150	50	150	300	100	75	250	80	0	0	0
1	1	0	102	–	0.87	0.11	0.01	10^{-10}	10^{-4}	10^{-2}	0	101	30	0	150	60	65	134	50	114	200	100	65	175	80	0	0.01	0
2	2	–	–	47	0.79	0.18	0.03	10^{-14}	10^{-5}	10^{-7}	0	101	34	0	150	56	65	134	50	114	200	82	65	175	69	0	0.01	0.03
3		–	–	34	0.95	0.03	0.01	10^{-17}	10^{-7}	10^{-9}	0	101	34	0	150	50	65	134	47	114	200	70	65	175	60	0	0.01	0.05
4		–	–	37	0.97	0.02	0.01	10^{-19}	10^{-8}	10^{-12}	0	101	34	0	150	48	65	134	46	114	200	64	65	175	56	0	0.01	0.07
5		0	70	–	1.0	10^{-3}	10^{-4}	10^{-25}	10^{-11}	10^{-13}	0	93	34	0	124	48	57	118	46	91	116	64	57	140	56	0	0.01	0.07
6		–	–	63	0.94	0.06	10^{-3}	10^{-23}	10^{-9}	10^{-20}	0	93	38	0	124	50	57	118	48	91	156	64	57	140	57	0	0.01	0.11
7		–	–	63	0.65	0.34	0.01	10^{-22}	10^{-8}	10^{-27}	0	93	41	0	124	52	57	118	49	91	156	64	57	140	58	0	0.01	0.15
8		–	–	47	0.63	0.36	0.01	10^{-23}	10^{-8}	10^{-32}	0	93	41	0	124	52	57	118	49	91	156	62	57	140	57	0	0.01	0.18
9		–	–	48	0.59	0.40	0.01	10^{-23}	10^{-8}	10^{-37}	0	93	42	0	124	51	57	118	49	91	156	60	57	140	56	0	0.01	0.21
10		–	–	64	0.25	0.74	0.01	10^{-22}	10^{-8}	10^{-45}	0	93	44	0	124	52	57	118	50	91	156	60	57	140	56	0	0.02	0.25
11		117	265	–	10^{-4}	0.03	0.97	10^{-18}	10^{-5}	10^{-54}	0.01	128	44	0.02	159	52	63	147	50	96	184	60	63	171	56	0	0.01	0.25
12		–	–	83	10^{-5}	0.12	0.01	0.24	0.64	10^{-15}	0	100	41	0	200	68	75	150	57	150	300	94	75	250	81	0	0.02	0.05
13		–	–	58	10^{-5}	0.45	0.06	10^{-3}	0.48	10^{-22}	0	100	43	0	200	65	75	150	57	150	300	85	75	250	75	0	0.01	0.09
14		–	–	62	10^{-6}	0.60	0.07	10^{-3}	0.33	10^{-30}	0	100	46	0	200	65	75	150	58	150	300	81	75	250	73	0	0.02	0.13
15		106	200	–	10^{-10}	10^{-4}	0.17	10^{-4}	0.82	10^{-40}	0.01	134	46	0.02	200	65	79	167	58	139	250	81	79	224	73	0	0.01	0.13

DISTRIBUTION LIST

1. Research Office (Code 09) 1
   Naval Postgraduate School
   Monterey, CA 93943-5000

2. Dudley Knox Library (Code 013) 2
   Naval Postgraduate School
   Monterey, CA 93943-5002

3. Defense Technical Information Center 2
   8725 John J. Kingman Rd., STE 0944
   Ft. Belvoir, VA 22060-6218

4. Therese Bilodeau (Editorial Assistant) 1
   Dept of Operations Research
   Naval Postgraduate School
   Monterey, CA 93943-5000

5. Prof. Donald P. Gaver (Code OR/Gv) 2
   Dept of Operations Research
   Naval Postgraduate School
   Monterey, CA 93943-5000

6. Prof. Patricia A. Jacobs (Code OR/Jc) 2
   Dept of Operations Research
   Naval Postgraduate School
   Monterey, CA 93943-5000

7. Dr. Mark A. Youngren 2
   4618 Duncan Drive
   Annandale, VA 22003

8. Dr. Samuel H. Parry 2
   245 South Joan Lane
   Gilbert, AZ 85296

9. Prof. Wayne Hughes (Code OR/Hl) 1
   Dept of Operations Research
   Naval Postgraduate School
   Monterey, CA 93943-5000

10. Prof. Thomas Lucas (Code OR/Lt) 1
    Dept of Operations Research
    Naval Postgraduate School
    Monterey, CA 93943-5000

11. Prof. Steven Pilnick (Code OR/Ps) 1
    Dept of Operations Research
    Naval Postgraduate School
    Monterey, CA 93943-5000

12. Prof. Dan Boger 1
    Chair, C3I Academic Group
    Naval Postgraduate School
    Monterey, CA 93943-5000

13. Prof. Phil DePoy.. 1
    Institute for Joint Warfare Analysis
    Naval Postgraduate School
    Monterey, CA 93943-5000

14. Prof. Gordon Schacher 1
    Physics Dept.
    Naval Postgraduate School
    Monterey, CA 93943-5000

15. LTC Jeffrey Appleget.. 1
    Director, TRAC-Monterey
    PO Box 8692
    Naval Postgraduate School
    Monterey, CA 93943-0962

16. Dr. Michael P. Bailey 1
    Principal Analyst, Modeling & Simulation
    MCCDC
    3300 Russell Road
    Quantico, VA 22134-5130

17. Mr. Michael Bauman.. 1
    Director, USA Training & Doctrine Command Analysis Center (TRAC)
    Fort Leavenworth, KS 66027

18. Dr. Alfred G. Brandstein 1
    MCCDC
    Studies and Analysis Division
    3093 Upshur Avenue
    Quantico, VA 22134-5130

19. Mr. Peter Byrne 1
    Joint Staff, J8
    Warfighting Analysis Division
    The Pentagon
    Washington, DC 20318-8000

20. Center for Naval Analyses 1
    4401 Ford Avenue
    Alexandria, VA 22302-0268

21. Mr. Yee Kah Chee.. 1
    Senior Manager (Operations Research)
    DSO National Laboratories
    20, Science Park Drive
    Singapore 118230
    REPUBLIC OF SINGAPORE

22. Mr. Wm. P. Clay 1
Director, USAMSAA
Attn: AMXSY-CA
APG, MD 21005-5071

23. Dr. Bruce W. Fowler 1
USAAMCOM
ATTN: ANSAM-RD-AS
Redstone Arsenal, AL 35898-5242

24. Dr. Arthur Fries 1
Institute for Defense Analysis
1800 North Beauregard
Alexandria, VA 22311

25. Dr. John T. Hanley, Jr. 1
Special Assistant to Commander-in-Chief
U.S. Pacific Command
Camp H M Smith, HI 96861

26. Mr. Walter W. Hollis 1
Deputy Under Secretary of the Army (OR)
ATTN: SAUS (OR)
The Pentagon, Room 2E660
Washington, DC 20310-0102

27. Prof. Reiner Huber 1
University BW Muenchen
Werner-Heisenberg Weg 39
Neubiberg D-85577
GERMANY

28. Dr. Beverly Knapp 1
US Army Research Laboratory-Ft. Huachuca Field Element
ATTN: AMSRL-HR-MY
Greely Hall (Bldg 61801) Rm. 2631
Ft. Huachuca, AZ 85613-5000

29. Dr. Moshe Kress 1
CEMA
P.O.B. 2250 (TI)
Haifa 31021
ISRAEL

30. COL R.S. Miller 1
US Army (Ret.)
Institute for Defense Analysis
1800 North Beauregard
Alexandria, VA 22311

31. Ms. Janet Morrow 1
Head, Modeling and Simulation
National Ground Intelligence Center
220 Seventh Street, NE
Charlottesville, VA 22902-5396

32. Mr. H. Kent Pickett 1
    Director, Modeling and Research Directorate
    TRAC Fort Leavenworth
    Fort Leavenworth, KS 66027

33. Mr. Vincent D. Roske, Jr. 1
    The Joint Staff, J8
    The Pentagon
    Washington, DC 20318-8000

34. Dr. Stuart Starr 1
    The MITRE Corporation
    7525 Colshire Drive, MSZ521
    McLean, VA 22102

35. Dr. William Stevens 1
    Metron
    512 Via de la Valle
    Suite 301
    Solana Beach, CA 92075

36. Mr. E.B. Vandiver III 1
    U.S. Army Concept Analysis Agency
    1820 Woodmont Avenue
    Bethesda, MD 20814-2979

37. Mr. Robert Wood 1
    Director, Center for Naval Warfare Studies
    Luce Hall
    Naval War College
    Newport, RI 02841

38. Mr. Yeo Siok Khoon 1
    DSO National Laboratories
    Operations Research
    20, Science Park Drive
    Singapore 118230
    REPUBLIC OF SINGAPORE